ANALYSIS OF TRANSCRIPTIONAL AND EPIGENETIC REGULATION IN *DROSOPHILA MELANOGASTER* MALE GERM CELLS DIFFERENTIATION

By

Cindy Lim

A dissertation submitted to the Johns Hopkins University in conformity with the requirement for the degree of Doctor of Philosophy

Baltimore, Maryland

October 1st 2013

Abstract

Understanding molecular mechanisms of stem cell maintenance and differentiation is key to effectively advance the use of stem cells in medicine. The *Drosophila* testis offers a particular insight in the studies of stem cell biology with its anatomically well-characterized niche structure. Moreover, *Drosophila* is a genetically tractable organism, which allows genetic manipulation and a multitude of tools generated by the scientific community to assess gene function with precise spatial and temporal control. Thus, Drosophila testis is an ideal model system in my quest to investigate the transcriptional and epigenetic regulation role in stem cell systems.

Asymmetric cell division is used in many stem cell systems to maintain a balanced population of stem cells and differentiating daughter cells. To understand systematically the mechanism of proper GSC differentiation, we analyzed transcription profiles of normally developing germ cells at discrete differentiation stages. We have developed a strategy to isolate a single cyst of germ cells (encapsulated by two somatic cells) from wild-type testis of *Drosophila* and successfully used these single germ cell cysts as the starting material for transcriptome profiling using RNA-sequencing technology. Together, the data from germ cell cysts at every stage of differentiation delineates a high resolution transcriptional profile of the entire germline lineage. Our analysis has resulted in several exciting findings with profound implications in our understanding of GSC differentiation.

It has been shown that differentiating progenitor cells at transit-amplifying stage can dedifferentiate and repopulate the niche. However, the molecular mechanism governing the progenitor cell dedifferentiation pathway is mostly unknown. Understanding how these progenitor cells undergo dedifferentiation will be critical to illuminate the processes of tissue homeostasis and injury repair. In our transcriptome screen, I identified a novel gene called

slamdance (sda) which is expressed specifically in the niche. The *sda* gene encodes an aminopeptidase, an enzyme whose function has not been previously shown in any stem cell system. We found that loss of function of *sda* leads to dramatic abnormalities in the niche, including deterioration of the niche architecture and the loss of stem cells. We determined that the loss of GSCs in the *sda* mutant is caused by defects in both dedifferentiation of progenitor cells and cadherin-dependent maintenance of the bonafide GSCs established in embryogenesis. Further studies showed that Sda is both necessary and sufficient to promote dedifferentiation of progenitor cells through the activity of its catalytic domain. Therefore, our findings provide *in vivo* evidence that a novel niche specific aminopeptidase promotes dedifferentiation to repopulate the stem cell niche under both physiological conditions and upon forced depletion of stem cells using genetic manipulation. Our results have advanced the understanding of how a niche-specific peptidase influences "differentiation versus dedifferentiation" decision of progenitor cells in an endogenous stem cell system.

After detailed analysis of transcriptional regulation which leads to the discovery of a novel dedifferentiation factor, we were interested in the role of epigenetic regulation in *Drosophila* stem cell system. Epigenetic modifications have been shown to contribute to maintaining stem cell identity, it has been a long-standing question whether and how stem cells retain their epigenetic information. To examine this question, we developed a novel switchable dual-color method to differentially label pre-existing and newly synthesized histones. We showed that pre-existing histone H3 is selectively retained in the male germline stem cells (GSCs) during asymmetric divisions. This asymmetric histone inheritance was observed in GSCs but not in symmetrically dividing progenitor cells. Furthermore, genetically manipulated GSCs that divide symmetrically lose this asymmetric histone inheritance. Lastly, we observed that in contrast to the canonical histones, the histone variant H3.3 does not exhibit this asymmetric pattern during GSCs division. In conclusion, our studies provided the first evidence that stem cells retain pre-existing

canonical histone during asymmetric cell divisions *in vivo*, which may contribute to maintain stem cell epigenetic memory and identity.

In conclusion, from our detailed analysis of the niche and germ cell differentiation lineage, we have discovered previously unknown epigenetic regulation in the niche and novel factors in the differentiation and de-differentiation process of germ cells which open doors to new exciting research.

Advisor: Xin Chen

Second Reader: Nicholas Ingolia

Acknowledgement

I would like to thank first and foremost my advisor, Dr. Xin Chen, for her support during my graduate studies. Xin is one of the smartest people I know. I learned curiosity, persistence, how to love science (even during the hard times), and the importance of time and resource management. More importantly, she is one of the few great mentors who encourages and expects independent thinking.

I am also blessed with having the most supportive thesis committee members: Dr. Allan Spradling, Dr. Mark Van Doren, and Dr. Nicholas Ingolia whose brilliant suggestions and invaluable inputs to the direction of my thesis project have always been dead-on and very insightful. I also have to thank all my mentors for their suggestions and candid feedback about making difficult career and life choices.

I will forever be thankful for all the people I have met during my graduate years. Especially members of Xin's lab who have become my surrogate family; their support was absolutely critical scientifically and personally. I will forever be grateful to Dr. Suk Ho Eun for all his scientific discussion & knowledge; Lama Tarayah for all the suggestions and discussions; Lijuan Feng for maintaining my flies and positive attitude to science; Vuong Tran, my dear co-author on more than one paper, for the endless coffee and ice-cream sessions that kept me sane (and awake) during graduate school; Dr. Zhen Shi, Dr. Qiang Gan, Dr. Jing Xie, Dr. Selena Kramer, and Dr. Dong Fen for their generosity and support.

I also wanted to thank my friends (too many to list here but you know who you are!) for providing the support and friendship that I needed. I would like to especially thank Thalia Farazi, Antonia Thomas, Diva Jonatan, and Lama Tarayrah for proof-reading this thesis so promptly and adhering to the schedule. Without these friends, my graduate school experience would have been wholly different and definitely less fun.

I especially thank my mother, father, sister, and brother whom I love so much. Without my family's unconditional love and support, I would never have made it this far.

The best outcome from my time in Baltimore, aside from scientific training, has been the privilege of meeting all the new faces and personalities. I feel blessed to have met so many genuinely kind people who were always willing to help each other. I apologize if I left out anyone. However, I am grateful for the support and help from everyone I have met during my years at Johns Hopkins University.

My last thanks go to Mark Osborn who has been a true and great supporter. His faith in me and my intellect instilled the confidence I needed to persevere and finish my last year. I hope we both can keep learning about life and strengthened our commitment and determination to each other while we live our life to the fullest.

Table of Contents

Index of Tables and Figures

Chapter 1: Introduction

Chapter 2: High resolution transcriptome analysis of *Drosophila* male germline stem cells lineage

Chapter 1:

Introduction

Summary

Stem cells, with their unique dual potential to self-renew and give rise to differentiating cells, play an essential role in development, tissue homeostasis, and tissue regeneration after an injury. Understanding the biology of stem cells and characterizing factors influencing the decision between self-renewal and differentiation are critical to effectively exploit them for medical therapies.

Adult stem cells normally reside in a micro-environment called niche. The decision to differentiate or to stay in the niche is influenced by a multitude of extrinsic and intrinsic factors. *Drosophila* spermatogenesis has become a model system to study mechanisms that regulate adult stem cell maintenance, proliferation, and differentiation since its niche is well-characterized anatomically and the differentiation process is molecularly characterized. Moreover, *Drosophila* is a genetically tractable model organism which offers molecular and genetic tools readily available from the community.

The *Drosophila* germline stem cells (GSCs) undergo a dramatic cellular differentiation to become distinctive mature sperm. This differentiation process has been shown to occur with dynamic transcriptional changes. In this thesis, I seek to understand how transcriptional and epigenetic regulation function in germline stem cell maintenance and its proper differentiation. This introduction chapter is aimed to summarize the current literature on transcriptional and epigenetic regulation in the Drosophila GSC system. I will begin by introducing the *Drosophila* spermatogenesis, and proceed to summarize the current knowledge of transcriptional and epigenetic regulation in niche and consecutive differentiating germ cells at mitotic and meiotic differentiation stages.

Drosophila spermatogenesis

The Drosophila testis is a long tubular structure. Anatomically, the stem cells niche is found at the apical tip of this organ and a linear distribution of progressively differentiated germ cells are found toward the basal end of the tube (Figure 1.1). A cluster of post-mitotic cells, called the hub, anchors both GSCs and Cyst Stem Cells (CySCs) to form a specialized micro-environment called the niche (Figure 1.2). The GSCs divide asymmetrically to self-renew and produce differentiating daughter cells termed gonialblasts (GBs) which are displaced away from the niche and start the differentiation program. Each GB undergoes four rounds of mitosis as a transit amplifying progenitor cell (spermatogonia cell) whose cytoplasm remains interconnected due to incomplete cytokinesis. The sixteen spermatogonia cells then synchronously undergo pre-meiotic S-phase before entering an elongated G2 phase called spermatocyte stage. In spermatocytes, the germ cells grow 25 times in volume and turn on a robust transcriptional program to express genes required for meiotic divisions and terminal differentiation (Fuller, 1998).

CySCs, two of which encapsulate each GSC, also divide asymmetrically (Cheng et al., 2011) resulting in one new CySC and cyst cell. In this way, the differentiating cyst consists of a gonialblast and two cyst cells. These two cyst cells will enclose synchronously dividing and differentiating germ cells until the individualization stage when the spermatid is separated and released into the seminal vesicle. This close physical association between cyst and germ cells throughout their differentiation progress suggests that these two cell types act cooperatively to advance their differentiation process.

Epigenetic regulation in the stem cells niche

In most species, adult stem cells reside in a specialized micro-environment termed niche. While the niche environment in many stem cell systems in other organisms has yet to be identified, in the fruit fly *Drosophila melanogaster*, the niche system for testis has been well-characterized. In this niche, the GSCs associate with two types of somatic cells: hub cells at the apical tip of the testis and CySCs. The niche provides a polarized environment where cell-cell adhesion and niche-to-stem cells close range signaling works in concert to maintain stem cell identity (Losick et al., 2011).

The maintenance of chromatin structure and gene expression can be intrinsically determined with asymmetrically segregating factors that support differentiation or "stem-ness". One such factor is epigenetic regulation that changes chromatin structure and gene expression without changing DNA sequences. Recently, more evidence that epigenetic mechanisms may play a role in regulating GSCs activity has been surfacing (Eun et al., 2010). The two major classes of chromatin regulators, ATP-dependent chromatin remodelers and histone modifying enzymes, have been identified to have a functional role in the male GSCs niche. For example, the Nucleosome Remodeling Factor (NURF) complex has been shown to positively regulate JAK-STAT signaling required for GSC and CySCs maintenance in the niche (Cherry and Matunis, 2010). NURF may promote the transcription of JAK-STAT activators or repress JAK-STAT inhibitors to regulate this pathway. Another example is a histone modifying enzyme which deubiquitinates mono-ubiquinated H2B, encoded by *scrawny (scny*), postulated to repress the transcription of differentiating genes to maintain GSCs; loss of Scny results in the loss of GSCs from the niche (Buszczak et al., 2009). Another chromatin factor called *no child left behind (nclb)*

functions male-specifically to maintain GSCs (Casper et al., 2011). Other examples include RNA binding proteins such as *Musashi* (Siddall et al., 2006), *Held-out-wings(Monk et al., 2010)*, and *IGF-II messenger RNA binding protein (Toledano et al., 2012)* which have been shown to be required in GSC maintenance. These examples highlight the importance of epigenetic regulation in the niche system. Recently, evidence has been emerging to support the likely hypothesis that epigenetic factors can cross talk with the signaling pathways regulating stem cells. One such example in the male GSCs system is demethylase dUTX which has been shown to play a role in regulating the JAK-STAT pathway in the niche (Tarayrah et al., 2013).

Molecular characteristic of stem cells niche

In the previous section, I have discussed the epigenetic regulation in the stem cell niche and the potential cross-talk it has with the signaling molecules that regulate the niche. In this section, I will focus on discussing the signaling pathways that have been shown to influence the niche micro-environment and thus stem cells maintenance. I will particularly focus on the pathways that regulate transcription in the stem cells niche.

The two major signaling pathways in the male GSCs niche are the Bone Morphogenetic Protein (BMP) pathway and Janus Kinase Signal Transducer and Activator of Transcription (JAK-STAT) pathway. In both pathways, the ligands [*Glass Bottom Boat (GBB)* and *Decapentaplegic (DPP*) for BMP; *Unpaired (Upd*) for JAK-STAT] are expressed and secreted from the niche to activate their receptors in the stem cells [*Saxophone (Sax), Thick veins (Tkv*), and *Punt* for BMP and *Domeless* for JAK-STAT] and promote the eventual translocation of downstream transcription factors [*Mothers against Dpp (Mad)* for BMP and *Stat92E* for JAK-STAT] which initiate the cascade of gene expression.

Although the direct target for both pathways has been identified in other *Drosophila* systems (Chen and McKearin, 2003a; Moustakas and Heldin, 2009; Rushlow et al., 2001; Song et

al., 2004), the direct targets in male GSCs system have yet to be identified. However, a downstream target for BMP may be known. In GSCs that are unable to respond to BMP pathway, a differentiation gene called *bag of marbles (bam)* is ectopically expressed and causes the GSCs to undergo premature differentiation and leave the niche (Kawase et al., 2004; Shivdasani and Ingham, 2003). Therefore, it implies that normally the BMP pathway represses *bam* transcription in male GSCs, a phenomenon which has been shown in the female GSCs (Chen and McKearin, 2003a).

The second signaling pathway in the niche, JAK-STAT pathway has been shown to play a major role in maintaining GSC-hub cell adhesion (Leatherman and Dinardo, 2010) which suggests that cell-cell adhesion molecules such as cadherin molecules are potential downstream target of the JAK-STAT pathway. A microarray study to identify gene expression changes in response to hyper-activated Stat discovered mostly genes expressed in CySC instead of GSCs (Terry et al., 2006). This suggests that active Stat signaling in somatic cells may predominate and is required for maintaining GSCs in the niche. Consistently, a recent study showed that ectopic expression in cyst cells of Stat92E target genes *Zinc finger homeodomain protein 1 (Zfh1)* or *Chronologically inappropriate morphogenesis (Chinmo),* which are both transcription factors, resulted in GSCs self-renewal outside of the niche(Flaherty et al., 2010; Leatherman and Dinardo, 2008). Moreover, Zfh1 has also been implicated in guiding GSCs self-renewal by possibly activating BMP signaling in the CySCs, providing a crosstalk between these two signaling pathways in the niche (Leatherman and Dinardo, 2008; Leatherman and Dinardo, 2010). Lastly, Suppressor of cytokine signaling 36E (Socs36E) which is an antagonist and direct target of the JAK-STAT pathways has been shown to play a role in maintaining a balanced ratio of CySCs and GSCs in the niche (Issigonis et al., 2009).

Dedifferentiation of mitotic cells

The paradigm of GSCs maintenance in the testis has been studied as solely a process of retaining GSCs in the niche; either by ensuring that stem cells adhere to the niche or repress differentiation so GSCs do not leave the niche. However, studies have shown that differentiated germ cells can dedifferentiate to re-occupy the stem cell niche and become GSC-like cells (Brawley and Matunis, 2004; Cheng et al., 2008; Sheng et al., 2009; Wallenfang et al., 2006; Wong and Jones, 2012). This phenomenon has been shown to also occur in the female GSC system (Kai and Spradling, 2004). These dedifferentiated germ cells have been shown to have distinct cellular features,such as misoriented centrosome, compared to *bona fide* GSCs but are still shown to undergo asymmetric cell division just like the GSCs (Cheng et al., 2008), which suggests that these dedifferentiated cells can still respond properly to the signals in the niche.

It has also been shown that meiotic cells called spermatocytes are unable to dedifferentiate to become GSCs (Sheng et al., 2009). This ability of spermatogonia to dedifferentiate suggests that at the spermatogonia stage, germ cells still have permissive chromatin landscape and plastic transcriptional profile to revert back to GSC-like cells. Moreover, this means that the decision to become spermatocytes involves irreversible commitment and changes in cellular identity. Finally, this irreversible change has also been observed in the female germ cells. It was reported that 4- to 8- transit amplifying cells can dedifferentiate in ovarian niche (Kai and Spradling, 2004), suggesting that dedifferentiation may be a more general mechanism in maintaining stem cells niche's homeostasis. Lastly, this dedifferentiation process is conserved in mouse testis with many cellular commonalities with *Drosophila* spermatogonial dedifferentiation (Barroca et al., 2009; Nakagawa et al., 2010).

Molecular characteristics of mitotic cells

Gonialblasts in Drosophila melanogaster undergo exactly four rounds of mitosis to give rise to 16-cell spermatogonia cysts. This exact round of mitosis has been shown to be regulated by the *bam* gene (Insco et al., 2009), which encodes a differentiation factor expressed in 4- to 16-cells spermatogonia with a peak expression in 8-cell spermatogonia (Gonczy et al., 1997). Bam protein accumulates to a threshold level that is required for spermatogonia to exit mitosis and enter meiosis (Insco et al., 2009). As the threshold is never reached in the absence of Bam, *bam* mutant testes are filled with continuously dividing spermatogonia cells, which provide a relatively homogenous population of spermatogonia without meiotic spermatocytes for genome-wide studies to obtain the transcriptional profile or chromatin status of mitotic germ cells.

A detailed analysis comparing the transcriptome of *bam* and wildtype tissues containing spermatocytes showed an enrichment of chromatin remodeling factors and histone modifying enzymes in *bam* testes (Gan et al., 2010b). Furthermore, Chromatin Immunoprecipitation followed by high-throughput sequencing (ChIP-Seq) in *bam* mutant testes revealed monovalent occupancy of either H3K27me3 repressive mark or lack of H3K27me3 and H3K4me3 activating mark on differentiation genes in spermatogonia enriched testes (Gan et al., 2010a). This is distinctive from the chromatin landscape of differentiation genes in Embryonic Stem (ES) cells that have bivalency of active and repressive mark which are thought to be the hallmark of chromatin landscape of differentiating genes in "soon to be" differentiating cells (Bernstein et al., 2006; Buszczak and Spradling, 2006; Guenther et al., 2007). Moreover, a recent study reported that an epigenetic reader encoded by *Plant Homeodomain Finger 7 (PHF7*) expressed specifically in GSCs and spermatogonia cells is required in GSCs and for proper spermatogonial differentiation(Yang et al., 2012). This study indicates that differentiation genes are still subject to dramatic changes at their chromatin regions in order to achieve robust transcription in

spermatocytes and there are yet to be discovered epigenetic regulators in this mitotic to meiotic switch.

The switch from mitosis to meiosis is very critical in the spermatogenesis process. An early transition to meiosis will lead to fewer germ cells and decreased fertility while failure to transition to meiosis will lead to germline tumors as the previously described *bam* mutant testes. An important component of this switch in testis involves the Epidermal growth factor (Egf) pathway. The Egf pathways consist of the ligand *Spitz* which is processed by a trans-membrane protease called *Stet* in germ cells (Schulz et al., 2002). Activated *Spitz* has been shown to act on *Egf receptors (Egfr)* expressed in somatic cyst cells (Kiger et al., 2000). It then acts through the *Guanine nucleotide Exhange Factor (EGF)* called *Vav* to activate *Rac-type small GTPases* which are fine-tuned in their regulation by antagonizing *Rho-type small GTPases* (Sarkar et al., 2007). Egfr signaling acts in cyst cells to restrict GSC self-renewal and allow proper GSC-to-gonialblast switch in the niche, allowing proper spermatogonial proliferation (Tran et al., 2000). It also promotes the switch from mitotic spermatogonia to meiotic spermatocytes, allowing proper germ cell differentiation.

Mutations in the serine/threonine kinase signal transducer encoded by *raf* have been shown to result in a similar phenotype to the *egfr* mutant, suggesting that the receptor tyrosine kinase (RTK) pathway may be required in cyst cells for proper spermatogonial differentiation. The direct target for both Egf and Raf pathways has yet to be identified. However, since compromised Egf and Raf signaling both lead to defects in germline-to-somatic cells interaction and also over-proliferation of spermatogonial cells, it is possible that the target genes regulate proper encapsulation of germ cells by the cyst cells which have been shown previously to have similar over-proliferation phenotype (Kai and Spradling, 2004; Lin et al., 1996; Parrott et al., 2011; Schulz et al., 2002).

Lastly, a nuclear envelope component, Nucleoporin 98-86, has been implicated to regulate proper GSC-to-gonialblast and spermatogonia-to-spermatocytes transitions (Parrott et al., 2011). Nucleoporin 98-86 has been shown to function upstream of the BMP pathways, JAK-STAT pathways, and the Egfr signaling pathways which potentially implicates the nuclear structure in regulating the cellular differentiation in germline lineage.

Transcriptional regulation in meiotic cells

The transition from spermatogonia to spermatocyetes is an irreversible commitment to the differentiation pathway (Sheng et al., 2009), which is accompanied by a series of transcriptional, epigenetic, and morphological changes. After the proliferation stage, germ cells undergo the last S phase followed by a prolonged G2 phase that initiates their entrance into the spermatocyte stage. During this spermatocyte stage, the germ cells grow 25 times in volume and turn on a robust transcription program to activate genes required for spermatocyte maturation, as well as genes needed for meiotic division and the terminal differentiation process. Most of these genes transcribed at the spermatocyte stage are required for later meiotic divisions and terminal differentiation have been shown to be under translational repression until the exact time when their encoded protein is required (Lin et al., 1996).

The G2 to M transition of meiosis I requires Cyclin B, Boule and Twine, which are all transcribed at the spermatocyte stage (Ayyar et al., 2003; Eberhart et al., 1996; White-Cooper et al., 2000). Boule, an RNA binding protein, translocates from the nucleus to the cytoplasm to trigger the G2/M transition by allowing the translation of Twine, homolog of Cdc 25 (Jiang and White-Cooper, 2003). At this time, Cyclin B protein also escapes its translational repression and accumulates in the cytoplasm of spermatocytes (Eberhart et al., 1996). Since both *boule* and *twine* mutant testis undergo spermatid differentiation independently of the meiotic cell cycle

progression (Jiang et al., 2007; White-Cooper et al., 2000), this suggests that these two processes can be uncoupled and independent from each other.

However, two classes of genes expressed in early spermatocytes show a high degree of coordination between meiotic divisions and spermatid differentiation (Perezgasga et al., 2004). Mutation in any of these genes resulted in meiosis arrest and block of spermatid differentiation, resulting in testes filled with immature spermatocytes. These "meiotic arrest" genes are classified into "aly-class" and "can-class" based on their morphological differences of the mutant spermatocytes (Eberhart et al., 1996; Perezgasga et al., 2004) and distinct target genes (Beall et al., 2007; Chen et al., 2005; Chen et al., 2011; Cler et al., 2009; Eberhart et al., 1996; Hiller et al., 2004; Hiller et al., 2001; Perezgasga et al., 2004; Wang et al., 2008; Wang and Mann, 2003; White-Cooper, 2009). For example, transcription of meiotic cell cycle genes such as *Cyclin B*, *boule*, and *twine* has been shown to rely on *aly*-class but not *can*-class genes (Eberhart et al., 1996). However, Boule protein accumulation has been shown to depend on the *can*-class genes. It was postulated that these meiotic arrest genes regulate transcription or translation of meiotic cell cycle genes so that the meiotic cell cycles do not prematurely proceed until terminal differentiation genes have been robustly transcribed.

The five known aly genes are *always early (aly)*, *cookie monster (comr)*, *matotopetli (topi)*, *tombola (tomb)*, and *achintya/vismay (achi/vis)*. Except for *achi/vis*, all of the *aly*-class genes are expressed exclusively in the primary spermatocytes (Chen et al., 2005; Chen et al., 2011; Hiller et al., 2004; Hiller et al., 2001; Metcalf and Wassarman, 2007; Tora, 2002). Four of the *aly*-class proteins have putative DNA domains and are thought to regulate transcription of yet to be identified target genes by directly binding to DNA sequences. Comr contains a winged helix; Topi contains multiple Zinc finger motifs; Tomb has CXC domains; Achi/Vis, products of gene duplication, have homeodomains.

Immuno-affinity purification studies have revealed that Aly and Tomb protein are co-purified with Mip40 to form the testis meiotic arrest complex (tMAC) which also contains Topi, Comr, and CAF1 (Wang and Mann, 2003). A second form of tMAC contains Aly, Comr, and Achi/Vis (Tora, 2002). The tMAC complex resembles the MIP/dREAM complex in mammals and the SynMuy complexes in *C. elegans* (Beall et al., 2007; Chen et al., 2005; Chen et al., 2011; Eberhart et al., 1996; Hiller et al., 2004; Hiller et al., 2001; Wang and Mann, 2003). Based on the resemblance to these complexes, tMAC is thought to activate transcription. Moreover, *achi/vis* mutant phenotype can be rescued by expression of Achi/Vis fused to a strong transactivation domain (VP16), but not a repression domain (EnR) (Chen et al., 2011). Consistent with its predicted role in transcription activation, all tMAC subunit is co-localized with euchromatin in primary spermatocytes.

The *can*-class genes encode testis-specific homologues of the ubiquitously expressed subunits of the general transcription factor II D ($TF_{II}D$). $TF_{II}D$ is one of the general transcription factors that constitute RNA polymerase II pre-initiation complex with TATA-binding protein (TBP) and 13-14 other TBP-associated factors (TAFs) to coordinate the interaction between the polymerase machinery and promoter regions (Cler et al., 2009; Matangkasombut et al., 2004; Tora, 2002). The characterized *can*-class genes include *no hitter (nht, TAF4L),cannonball (can, TAF5L), meiosis I arrest (mia, TAF6L), spermatocyte arrest (sa, TAF8L),* and *ryan express (rye, TAF12L)*. Four of these can-class genes, namely Mia, Nht, Rye and Sa, have histone folding motifs for protein-protein interaction. Moreover, Nht and Rye are found to hetero-dimerize *in vitro* (Hiller et al., 2004). On the other hand, Can is a WD40-repeat-containing protein (Hiller et al., 2001).

As $TF_{II}D$, Can- class genes are thought to be required in transcriptional activation of terminal differentiation genes in a testis-specific manner (Hiller et al., 2004; Hiller et al., 2001). This would suggest that tTAF (testis-specific TBP (Transcription Binding Protein) Associated

Factor) would localize to euchromatin in spermatocyte nuclei. However, only a proportion of the total protein of each tTAF components associate with chromosomes in spermatocytes. The majority of the tTAF protein was found in a sub-compartment within the nucleolus (Chen et al., 2005; Metcalf and Wassarman, 2007). This sub-nucleolus compartment has been shown to also contain Polycomb (Pc) and other components of the Polycomb Repressive Complex 1 (PRC1) in spermatocytes. Further studies found that localization of PRC1 components to the spermatocyte nucleolus is coincident with tTAF expression and dependent on wild-type tTAF function (Chen et al., 2005), which suggest that tTAFs potentially act as de-repressors by sequestering PRC1 to counteract PcG-induced repression. However, recent studies found that in the absence of tTAF, removing PcG activity is not sufficient to turn on terminal differentiation genes (Chen et al., 2011), suggesting tTAFs are still required for transcription activating element. Consistently, tTAFs have been reported to potentially activate transcription of more than 1,000 genes, many of which are required for spermatid differentiation (Chen et al., 2011; White-Cooper et al., 1998). Another compelling evidence is that three of these tTAF-dependent genes are direct target genes by ChIP assay: *fuzzy onions* (*fzo*), required for mitochondrial fusion in early spermatids (Hales and Fuller, 1997), *mst87F*, a component of the sperm tail(Schafer et al., 1993) and *don juan* (*dj*), a sperm-specific DNA-binding protein localizes to mitochondria (Santel et al., 1998). Furthermore, ChIP analysis showed that at the promoter regions of these genes, the levels of the repressive H3K27me3 mark and paused Pol II are high in contrast to the low level of active H3K4me3 mark that are low in *can* and *aly* mutant testes (Chen et al., 2011), suggesting that tTAFs and tMAC may function by recruiting the Trithorax group complex (TrxG) that antagonize PcG by methylating H3K4 at their promoters (Chen et al., 2005).

Although the interaction between tMAC components (*aly*-class) and tTAFs (*can*-class) is yet to be understood fully, *aly* has been shown to be required for TAF8L binding to target gene promoters. Moreover, as Aly is also required for the proper nucleolar localization of several tTAFs and Pc in spermatocyte cells, it is highly likely that tMAC acts upstream of tTAFs (Chen

et al., 2011). Additionally, while Mip40 is co-immunoprecipitated with tMAC components, the loss of *mip40* results in similar phenotype to *can*-class mutants (Beall et al., 2007). These results are highly indicative that Mip40 may be the mediator of tMAC and tTAFs interaction.

Lastly, a recent study has identified two additional meiotic arrest genes which cannot be classified as either the *aly*-class or the *can*-class (Doggett et al., 2011; Moon et al., 2011). Wake-up-call (Wuc) is highly expressed and associated with chromatin and physical interact with Aly in a yeast-two-hybrid screen (Jiang et al., 2007). However, loss of *wuc* does not abolish expression of either meiotic cell cycle genes or spermatid differentiation genes (Doggett et al., 2011).The second gene, THOC5 is a part of the THO complex known to export mRNA from the nucleus to the cytoplasm. Loss of this gene led to meiotic arrest phenotype with disrupted nucleolar structure and tTAF localization (Moon et al., 2011). However, no mRNA export defects or decreased transcription of meiotic cell cycle genes or spermatid differentiation genes were detectable in *thoc5* mutant. In conclusion, the identification of *wuc* and *thoc5* mutants illuminates other meiotic arrest genes other than *aly*-class and *can*-class which demonstrate that there is still new information and knowledge to be learned about spermatocyte maturation.

Epigenetic memory to maintain cell fates

In eukaryotic nuclei, DNA is organized into nucleosomes by wrapping around histone octamers [2×(H3, H4, H2A, H2B)], which act as fundamental units to form the higher-order chromatin structure. The deposition of histone, as a potential mechanism of epigenetic memory regulation will be discussed in the next section of this introduction. Epigenetic mechanisms that alter chromatin structure while preserving primary DNA sequences contribute significantly to "cellular memory", which maintains a particular cell state through many cell divisions (Jacobs and van Lohuizen, 2002; Ringrose and Paro, 2004; Turner, 2002). It has been shown that two major epigenetic regulations could occur in two levels: 1. DNA structural level such as DNA

methylation; 2. DNA-associated protein level such as histones. There is more evidence supporting the hypothesis that extensive post-translational modifications of histones have profound impact on regulating gene expression (Berger, 2007; Fischle et al., 2003; Jenuwein and Allis, 2001; Schreiber and Bernstein, 2002; Turner, 2002). Thus, it is possible that stem cells have a chromatin structure distinct from their differentiated cells which maintain stem cell unique molecular characteristics (Eun et al., 2010; Jaenisch and Young, 2008; Jenuwein and Allis, 2001)

DNA methylation is one epigenetic regulation whose inheritance is best understood due to its semi-conservative propagation (Bonasio et al., 2010; Martin and Zhang, 2007). Additionally, certain histone modifications at constitutively active genes, such as the hyper-acetylation of H3 and H4, H3K4me2/3 and H3K79me2 marks, are maintained in mitotic cells when global transcription is shut off. This suggests that histone modifications could serve as molecular memory bookmarks to recapitulate transcriptional active domain after mitosis (Kouskouti and Talianidis, 2005; Valls et al., 2005). On the other hand, repressive histone modifications such as H3K9me3 have also been shown to remain associated with chromatin during mitosis (Fischle et al., 2005), which is probably important for faithful inheritance of heterochromatin structure (Irvine et al., 2006; Motamedi et al., 2004). However, it still remains unclear whether and how the histone modification patterns could be inherited or reestablished asymmetrically in stem cells and their differentiating daughter cells.

Histone depositions into the chromatin structure

As previously stated, histone marks can potentially be a major epigenetic memory carrier for stem cell state. As such, it is important to understand how histones are deposited into the chromatin structure as it can illuminate the mechanism of epigenetic memory with histone as the

carrier. Current literature has distinguished two ways as to how histones are deposited: replication-coupled and replication-independent.

The bulk of canonical histones are synthesized and incorporated during DNA replication when the entire genome duplicates. During this process, it is commonly accepted that H3 and H4 are incorporated as a tetramer, while H2A and H2B are incorporated as dimers (Annunziato et al., 1982; Jackson and Chalkley, 1981a; Jackson and Chalkley, 1981b; Russev and Hancock, 1981; Xu et al., 2010). Replication dependent histone deposition is a highly regulated process, which requires an orchestrated series of events such as the disruption and recycling of preexisting octamers and deposition of newly synthesized histones at the replication fork (Corpet and Almouzni, 2009). Abnormal nucleosome assembly and deposition has been shown to cause genome instability and increased sensitivity to DNA damage which eventually lead to tumorigenesis and other diseases. Thus, the process of histone deposition must be coordinated very efficiently and precisely to assemble chromatin right after the passage of the replication fork (Gasser et al., 1996; Guilbaud et al., 2011; Smith and Whitehouse, 2012).

Incorporation of newly synthesized histones is facilitated by chromatin remodeling complexes (Saha et al., 2006), histone chaperones (De Koning et al., 2007) and histone modifying enzymes (Corpet and Almouzni, 2009). Modifications such as acetylation or methylation to the N-termini of the histone have been shown to be critical in this deposition process (Ai and Parthun, 2004; Masumoto et al., 2005), by allowing histones to interact with histone chaperones (Chen et al., 2008; Li et al., 2008; Recht et al., 2006; Verreault et al., 1996) and the DNA replication machinery (Moggs et al., 2000; Shibahara and Stillman, 1999). Despite increased knowledge on incorporation of newly synthesized histones during DNA replication, our understanding on whether and how preexisting histones are recycled at replication forks is limited. It is yet unclear how reliable the transmission of histone modification information from the mother cell to the two daughter cells actually is. Moreover, how this information is distributed

during asymmetrical cell division so that each of the daughter cells has different cell fates remains unclear.

Increasing evidence indicates that histone variants also influence epigenetic inheritance (Henikoff et al., 2004a; Henikoff et al., 2004b). However, in contrast to canonical histones, deposition of the histone variants is replication-independent. One of the best characterized histone turnover during transcription is the replacement of canonical H3 with the H3.3 variant (Ahmad and Henikoff, 2002b; Schwartz and Ahmad, 2005). This replacement is more frequent at gene regulatory regions, actively transcribed coding sequences, and replication origins (Deal et al., 2010; Mito et al., 2007). This process requires histone chaperones such as HirA (Ray-Gallet et al., 2002), Daxx and Atrx (Drane et al., 2010; Goldberg et al., 2010), as well as chromatin remodeling complexes (Konev et al., 2007). It was postulated that histone variants such as H3.3 can potentially transmit either the active or repressive chromatin state as epigenetic memory to maintain similar gene expression during mitosis or meiosis (Szenker et al., 2011). Other histone variants have been shown to play distinct roles in a variety of different cellular processes (Kamakaka and Biggins, 2005), such as CENP-A, a H3 variant found specifically at the centromere region and deposited in a replication-independent manner (Ahmad and Henikoff, 2002a).

Conclusions

As shown in this chapter, there have been many recent discoveries in molecular characterization of the stem cell niche and its differentiation pathways which are particularly compelling since *Drosophila* and its mammalian counterparts have been shown repeatedly to share conserved signaling and regulation pathways. For example, Stat 3 has been shown to regulate differentiation in mouse testis (Oatley, 2010) while Smad proteins, the homologs of Drosophila Mad, have also been shown to regulate the progression of germ cell differentiation in adult mouse testis (Itman and Loveland, 2008). The plethora of knowledge about regulation and interactions between these pathways in the stem cell niche gleaned in *Drosophila* has shown that this endogenous stem cell system provides an ideal model system.

New technologies, including high-throughput mRNA sequencing (RNA-seq) and ChIP-seq, have opened the door to more exciting research previously hampered by the limitation of starting material. In my thesis work, I took advantage of the sensitivity of new technologies. I combined this highly sensitive genomic technique with isolation of pure cells at distinct stages of spermatogenesis to explore the transcriptional profiles with unprecedented spatio-temporal resolution.

Figures & Tables

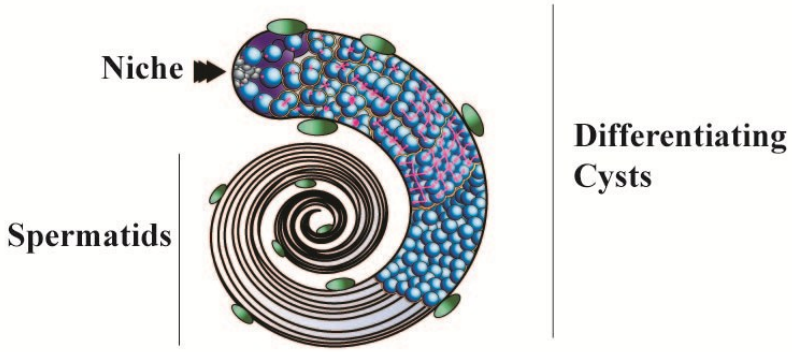


Figure 1-1 : Diagram of adult testis. The testis has a stem cell niche at the apical tip of the tissue. This tip region contain the undifferentiated germ cells (blue) while the remainder of the testis is filled with differentiated germ cells still connected through a structure called fusome (purple). The germ cells then undergo meiosis and terminal differentiation to yield an elongated specialized cell called the sperm.

Niche.

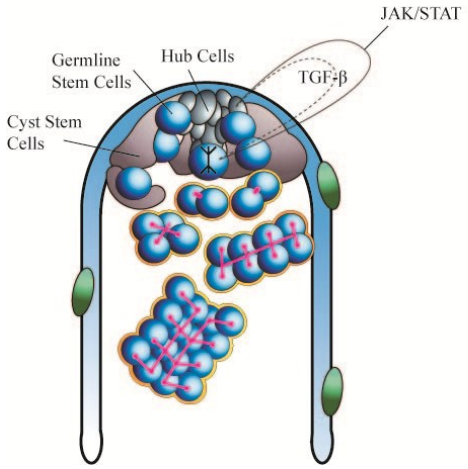


Figure 1-2 : Diagram of the testis niche. The testis niche contained the hub cells (grey), the germline stem cells-GSCs (blue) and the cyst stem cells- CySCs (pink). The niche are located at the tip of the testis. Asymmetric cell division yield self-renewed stem cell and differentiating daughter cells. The germ cells undero incomplete cytokinesis and the germ cells in the cyst are still interconnected through a fusome structure (purple). The GSCs receive JAK/STAT and TGFβ signals from the hub.

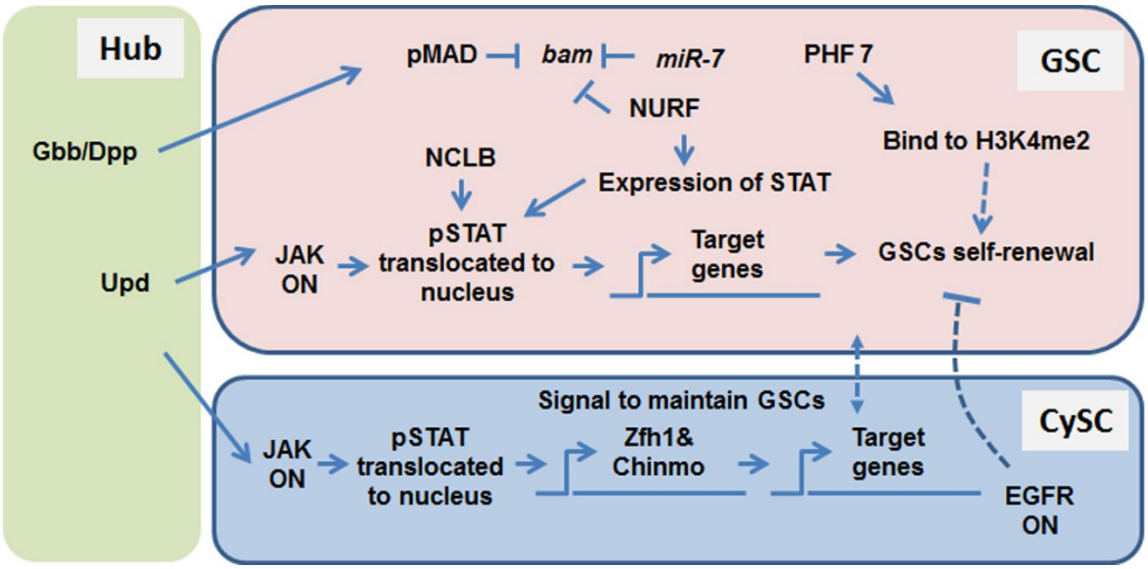


Figure 1-3 : : Summary of transcriptional regulation in stem cell niche. Hub cells are in green, GSC cells are in pink, CySCs are in blue. Solid lines denote direct regulation, dashed lines denote indirect regulation or lack of evidence for direct regulation.

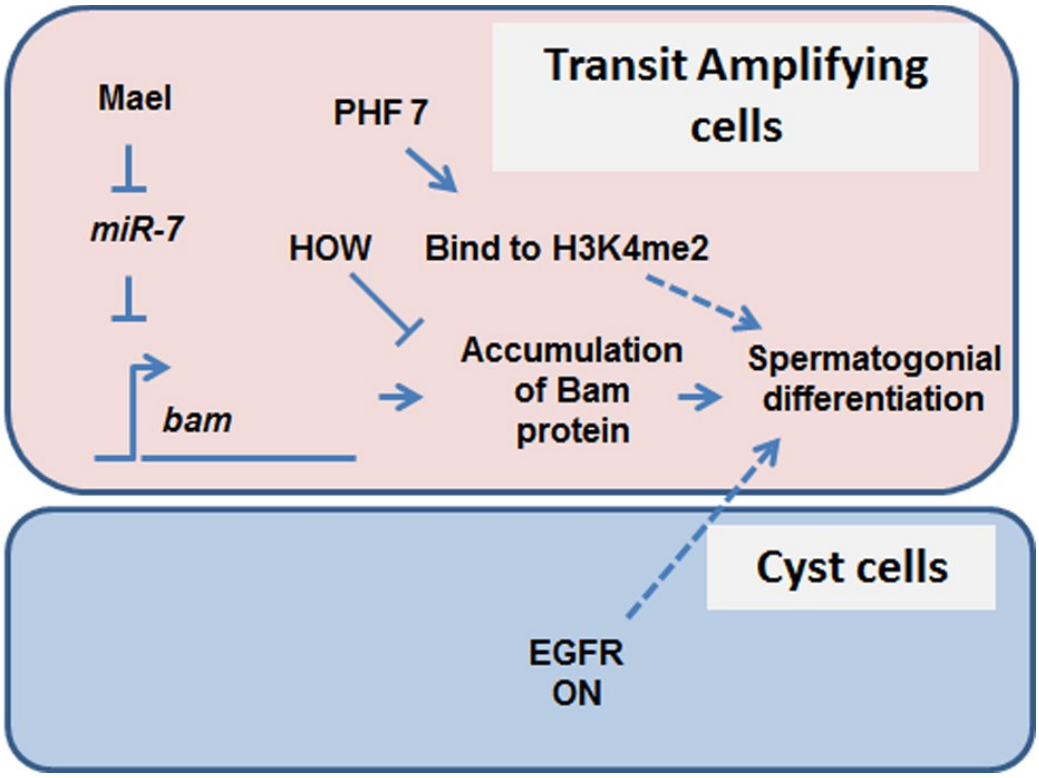


Figure 1-4: : Summary of transcriptional regulation in mitotic germ cells. Transit-amplifying germ cells are in pink, Cyst cells are in blue. Solid lines denote direct regulation, dashed lines denote indirect regulation or lack of evidence for direct regulation.

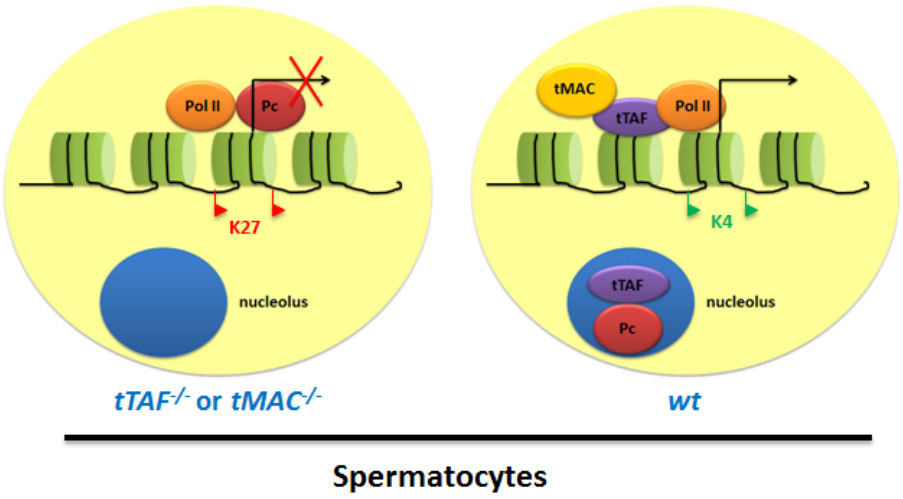


Figure 1-5: Summary of transcriptional regulation in meiotic and post-meiotic germ cells. A schematic diagram outlines potential chromatin state in spermatocytes mutant for tMAC or tTAF (left, analyzed with *can* or *aly* mutant testes) compared to mature wild-type (wt) spermatocytes (middle, analyzed with wt testes). K27: H3K27me3, K4:H3K4me3.

Cell Types	Stage	Gene name	Function(s)
Somatic Cells	Hub	Gbb/Dpp	Ligand to TGF-beta signaling
		Upd	Ligand to JAK-STAT signaling
	Cyst Stem Cells	Rac	GTPase, downstream of Egf pathway, inhibit GSCs self-renewal
		pSTAT	Transcription factors activating JAK-STAT target genes, sufficient for GSCs self-renewal
		Chinmo	Transcription factors, a JAK-STAT target genes, sufficient for GSCs self-renewal
		Zfh1	Transcription factors, a JAK-STAT target genes, sufficient for GSCs self-renewal
		Socs36E	JAK-STAT antagonist, maintain balance of CySCs and GSCs ratio in the niche
	Cyst Cells	Rac	GTPase, downstream of Egf pathway, promote spermatogonial differentiation
Germ Cells	GSCs	Nucleoporin98-86	nuclear envelope component, regulate proper GSC to GB transition
		pMAD	Transcriptional regulator of TGF-beta target genes
		pSTAT	Transcription factors activating JAK-STAT target genes, increase GSC-Hub adhesion
		Scny	deubiquitinating enzyme, target mono-ubiquitinated H2B, repress differentiation genes
		NCLB	male-specific chromatin factors, decrease expression or accumulation of STAT92E
		NURF	Chromatin remodeler, positively regulate JAK-STAT signaling
		PHF7	Epigenetic reader, bind to H3K4me2, required for GSCs maintenance
		miR-7	small RNA, bind to *bam* mRNA, down regulate *bam* expression
		Msi	RNA binding protein, required for stem cells maintenance
		HOW	RNA binding protein, required for GSCs proliferation
		Imp	RNA binding protein, stabilizes *upd* mRNA
	Spermatogonia	Bam	Differentiation factor, required for spermatogonia differentiation
		Mael	RNA binding protein, repress miR-7, upregulate *bam* expression
		PHF7	Epigenetic reader, bind to H3K4me2, required for proper spermatogonia differentiation
		HOW	RNA binding protein, required for spermatogonia proliferation
		Nucleoporin98-86	Nuclear envelope component, regulate

			proper spermatogonia to spermatocytes transition
	Spermatocytes	*aly*-class (tMACs) *aly*(*always early*)	Regulation of tTAFs nucleolar localization and binding to target genes
		achi+vis (*achintya + vismay*)	Transcriptional activation of meiotic cell cycle genes
		comr (*cookie monster*)	Meiotic division and differentiation of spermatocytes
		topi (*matotopetli*)	
		tomb (*tombola*)	
		mip40 (*Myb interacting protein 40kD*)	
		can-class (tTAFs) *can* (*cannonball*)	
		sa (*spermatocyte arrest*)	Transcriptional activation of terminal differentiation genes
		mia (*meiosis I arrest*)	Accumulation of Boule protein
		nht (*no hitter*)	Meiotic division and differentiation of spermatocytes
		rye (*ryan express*)	
		wuc (*wake up call*)	Meiotic division and differentiation of spermatocytes
		THO-compex *thoc5*	Meiotic division and differentiation of spermatocytes
		thoc 7	
		tho2	
		hpr1	
		NURF (Nucleosome Remodeling Factor)	Meiotic division and differentiation of spermatocytes
	Spermatids	Comets and Cups	

Table 1-1 : Factors required in transcriptional regulation of spermatogenenesis

Chapter 2:

High resolution transcriptome analysis of *Drosophila* male germline stem cells lineage

Summary

To effectively utilize stem cells in regenerative medicine, the molecular mechanisms underlying their maintenance and proper differentiation must be thoroughly understood. To achieve this goal, we used *Drosophila* male germline stem cell (GSC) lineage as a model adult stem cell system and systematically analyzed transcriptome of normally developing germ cells at discrete differentiation stages. We first developed a strategy to isolate a single cyst of germ cells (encapsulated by two somatic cells) at each stage from *wild-type* testes of *Drosophila*. We then applied these single germ cell cysts as the starting material for high-throughput mRNA sequencing (RNA-seq). Our data from every stage germ cell cysts delineates a high-resolution transcriptional profile in the entire male germline lineage and leads to multiple novel discoveries. (1) We found distinct transcriptional profiles between GSC niche sample and the immediate daughter cell of GSC, gonialblast (GB) cyst sample. In particular, Toll-pathway genes are highly expressed in the GSC niche sample. Furthermore, knockdown of Toll-signaling pathway components led to dramatic defects in the GSC niche. (2) We found that the GB cyst sample has enriched chromatin regulators, which may contribute to reset the chromatin structure for committing to differentiation. (3) We found that the transit-amplifying cells, known as spermatogonia, share a high degree of similarity in their overall transcriptomes, which explains their plasticity shown by their similar behavior during both dedifferentiation and differentiation processes. (4) We confirmed a dramatic transcriptome switch from mitotic spermatogonia to early meiotic spermatocytes, and found that many chromatin and transcriptional regulators show a bi-modal expression pattern. (5) We found that the maturation from early to late spermatocyte is accompanied by another substantial transcriptome change by turning on meiotic division and terminal differentiation genes, many of which are potential target genes of those regulators turned on in early spermatocytes. (6) Finally, we analyzed dosage compensation and found little compensation of X-chromosomal genes in germline cysts at each differentiation stage. In

summary, our single cyst-resolution, genome-wide transcriptional profile analyses provides a supreme and unprecedented data set to understand many interesting question in stem cell and germ cell biology fields.

Introduction

Maintenance of a multi-cellular organism during homeostasis and tissue repair requires active replenishment of aged or injured cells. Adult stem cells can fulfill this requirement due to their unique ability to both self-renew and give rise to differentiated special cell types (Betschinger and Knoblich, 2004; Clevers, 2005; Inaba and Yamashita, 2012; Morrison and Kimble, 2006). In many adult stem cell lineages, progenitor cells undergo a proliferative stage to expand before commitment for terminal differentiation. The switch from proliferation to differentiation must be tightly regulated and mis-regulation may lead to tumorigenesis or tissue dystrophy (Clarke and Fuller, 2006). On the other hand, progenitor cells remain plastic and can dedifferentiate in multiple stem cell lineages, which is critical in replenishing lost stem cells during aging or injury (Barroca et al., 2009; Boyle et al., 2007; Brawley and Matunis, 2004; Cheng et al., 2008; Friedmann-Morvinski et al., 2012; Kai and Spradling, 2004; Nakagawa et al., 2010; Schwitalla et al., 2013; Sheng et al., 2009; Talchai et al., 2012; Wallenfang et al., 2006; Wong and Jones, 2012). In order to be able to properly differentiate adult stem cells or progenitor cells *in vitro* and/or to promote dedifferentiation *in vivo* for regenerative medicine, researchers have to fully understand the molecular mechanisms underlying the normal differentiation program.

Drosophila spermatogenesis is gaining attention as a model system for the study of mechanisms regulating the maintenance and proliferation of stem cells, as well as proper differentiation of precursor cells (Brawley and Matunis, 2004; Kiger et al., 2001; Tulina and Matunis, 2001; Yamashita et al., 2003; Yamashita et al., 2007) (Fig.1A). In *Drosophila* testis, germline stem cells (GSCs) can be located precisely by their proximity to a group of post-mitotic cells called hub cells. A GSC divides asymmetrically to self-renew and gives rise to a gonialblast (GB), the daughter cell that initiates differentiation. GBs first go through a transit-amplifying stage, in which cells undergo exactly four rounds of mitosis. Once spermatogonial proliferation is

complete, germ cells enter the spermatocyte stage, in which each cell grows approximately 25-fold and initiates a robust gene expression program that enables meiotic division and spermatid differentiation [reviewed in (Davies and Fuller, 2009; Fuller, 1998; Lim et al., 2012)]. In parallel with GSC division, the cyst stem cell (CySC) surrounding each GSC also divides once. One CySC's daughter retains stem cell identity, while the other becomes a cyst cell, which encapsulates differentiating germ cells and never divides again. It has been demonstrated that CySCs and cyst cells communicate with accompanying germ cells *via* multiple signaling pathways for their critical decisions throughout spermatogenesis (Chen et al., 2013; Kiger et al., 2000; Leatherman and Dinardo, 2008; Leatherman and Dinardo, 2010; Lim and Fuller, 2012; Parrott et al., 2012; Sarkar et al., 2007; Schulz et al., 2004; Tran et al., 2000). The dramatic cellular differentiation during *Drosophila* spermatogenesis is accompanied by dynamic changes in gene expression, which is orchestrated by both extrinsic cues such as paracrine factors that initiate signaling pathways, as well as intrinsic factors such as chromatin regulators [reviewed in (Davies and Fuller, 2008; de Cuevas and Matunis, 2011; Lim et al., 2012)].

Previous studies have tried to dissect the transcription networks underlying germline stem cell differentiation by comparing gene expression profiles of mutant testes that accumulate germ cells at distinct stages of cellular differentiation to *wild-type* testes (Chen et al., 2011; Gan et al., 2010a; Terry et al., 2006). Although these approaches have provided useful information and led to functional studies using candidate gene approach, the information gleaned from intact tissues are limited by the inherently mixed population of cells present in the mutant background, and the difficulty in extrapolating results obtained from mutant backgrounds to normal germline stem cell regulation in *wild-type*. Here we use state-of-the-art cellular and molecular strategies to systematically study the gene expression of GSC lineage in every single recognizable differentiation stage. We are interested in the following questions: (1) Do GSCs and GBs, the two daughter cells derived from GSC asymmetric divisions, have similar or distinct transcriptional

profiles? (2) How does the transcriptome change in continuously proliferating spermatogonial cells? (3) Does the switch from mitosis to meiosis accompany with a transcriptome change that leads to a second transition of transcriptome change during spermatocyte maturation? (4) How do changes in the transcriptional program lead to cellular defects? The studies of single-cell transcriptome has provided a comprehensive dataset at a resolution that has never been achieved before, which yielded much-needed information on transcriptional status at each critical stage from an endogenous stem cell system.

Materials and Methods

Fly strains and tissue preparation

Fly stocks were grown in standard Bloomington medium at 25 °C. The upd-gal4 (a gift from T. Xie), bam-gal4 (a gift from Dennis McKearin), sa-gfp (Chen et al., 2005), dj-gfp (Bloomington Stock Center, BL5417), and UAS--tub-GFP (Bloomington Stock Center, BL 7373). Testis were dissected in S2 insect media (Invitrogen # 11720-034) + 10 % FBS (Invitrogen # 16000-044). Two micro-needles were used to tear the muscle sheath layer, causing the germ cell cysts to spill out into the media. The cysts were viewed using inverted microscope (Zeiss, Axiovert 40 CFL) and were picked using glass needle (Drummond, 1-000-0250) pulled in Sutter instrument P97 using the following conditions P=200; Heat= 669; Pull =30; Velocity=120; Time= 200. The picked cyst was transferred into PCR tubes by grounding the fine tip to the sides of the tubes.

There is an ambiguity in assigning stages to the isolated cysts due to technical limitation in our samples isolation. For example, spermatogonia samples can be distinguished by expression of bam-GFP marker. However, any 4 cells cyst isolated can potentially come from a broken cyst from 8S or 16S samples. Although it is a low possibility for 8S cysts to precisely "break" into 4 cells cysts or 16S cysts to "break" into precisely 4 cells cysts, we have tried to prevent this caveat by selecting for cysts that has "smooth" sphere appearance. Nonetheless, we decided to isolate two biological replicates for each of these "ambiguous samples", namely: GB, 4S, and 8S. Our results suggest that we isolated unbroken cysts since the biological samples are highly correlated with each other, specifically the GB, S4, and S8 samples pairs all have >0.9 Pearson's correlation coefficients. This high correlation between biological replicate also suggested the high reproducibility of single cyst isolation and sequencing procedures.

Transition from early to late spermatocytes consists of 3.5 days (~40%) in the spermatogenesis process which takes total of 10 days. It also involves almost 20 times growth in volume. We can accurately isolate EC by picking 16 cells GFP positive cells that has the same size as 16S GFP negative cells (16S) as these EC just recently become spermatocytes and began expressing sa-GFP marker. However, we do not have a good cell size comparison to distinguish LC samples from earlier spermatocytes stages. We decided to isolate 2 LC samples by visually choosing the largest Sa-GFP cells. Our results showed that the two LC samples we isolated (LC1 and LC2) deviate significantly from each other. This is likely due to the different spermatocytes maturation stage of LC1 and LC2. LC1 is more likely at earlier maturation stage compared to LC2 since it is more correlated with EC sample compared with LC2.

Library preparation for RNA-seq

The mRNA libraries were prepared according to methods described in (Kurimoto et al., 2006) with the modification listed in (Tang et al., 2009). In summary, we always process the isolated cyst immediately after isolation. First, we treated the cysts with mild detergent to permeabilize the plasma membrane while keeping the nucleus intact. Permeabilization allowed mRNA to be reverse transcribed using polydT attached to universal sequence. Poly A was added to the single strand DNA 3' end using terminal transferase. Second strand DNA was synthesized using poly dT attached to another universal primer sequence. Finally, the resulting library was amplified using both universal primers. To generate libraries for sequencing, ~300 ng dsDNA of each sample was fragmented by sonication using Bioruptor (Diagenode, UCD-200-TM-EX) under medium power output for 30 min in ice water. The resulting DNA fragments were analyzed by agarose gel to verify a ~100-300 bp size range. Sequencing libraries were prepared as the follows: end-repair (DNA end-repair kit from Epicenter, cat. no. ER0720); A-tailing (300 ng dsDNA, 5 µl Thermo buffer, 10 nmol dATP, 15 U of Taq polymerase, at 70 °C for 30 min); Solexa adaptor ligation (300 ng dsDNA, 4 µl DNA ligase buffer, 1 µl Solexa adaptor mixture, 3µl DNA ligase, at 70 °C overnight.); PCR (98 °C 10 s, 65 °C 30 s, 72 °C 30 s for 16 cycles; then additional 72 °C for 5 min) amplification with adaptor primers; and size selection (200-400 bp). Lastly, dsDNA library for each sample was used on Solexa 1G sequencer at a concentration of 10 ng per lane.

Gene annotation and gene expression calculation

We used EdgeR package (Robinson et al., 2010) to quantitatively measure the level of transcripts normalized across different samples which is termed corrected Read per Kilobase of gene model per Million mapped reads (cRPKM). cRPKM generated by edgeR algorithm

compensates for the difference in the sequence reads between samples, as well as the different number of samples within each stages compared. Thus, the provided cRPKM values are comparable between samples' libraries.

Read counts from biological replicates are fitted into negative binomial distribution model to obtain a dispersion estimate for each gene ("tagwise dispersion" in EdgeR), followed by differential expression inference. One single estimate of transcript concentration is evaluated by maximum likelihood method (qCML in EdgeR), for biological replicates. This single estimate is termed as "estimated transcription level".

Alignment of short reads onto Drosophila genome and assignment into annotated gene regions

36-nucleotide-long single-end short reads were collected from Illumina GAIII sequencer in Fastq format. There are **13** libraries of short reads from 10 distinct spermatogenesis stages, with three biological duplicates in GB, S4 and S8 stages as described in the previous section. BOWTIE software (ref, version 0.12.7) was utilized to align these short reads to Drosophila melanogaster genome sequence (Flybase dmel_r5.43, as of Jan 2012, ftp://ftp.flybase.net/releases/FB2012_01/dmel_r5.43/). Only those short reads that are uniquely mapped and at most two mismatches were retained for later analysis (later referred to as "aligned"). The detailed parameters for running BOWTIE are -a --phred33-quals -n 2 -e 70 -l 28 -m 1 --best –strata. Those filtered short reads were further searched for possible mapping onto exon junction regions due to splicing of pre-mRNA using TOPHAT software (version 1.3.3). The same criteria were performed as when running BOWTIE. The detailed running parameters are as followed: -g1 --butterfly-search -I 50000 --segment-length 15 --max-segment-intron 50000 -G. The annotation of exon structures are described in the following paragraph. Approximately only 1% of aligned reads are from splice junction regions. The reads aligned to the genome sequence

and the junction reads are combined and sifted for non-redundancy. At last, we concentrated on approximately two million non-redundant reads per stage sample.

We then assigned each read into gene regions. The annotation for protein coding genes, ribosomal RNAs, tRNAs, snoRNA, snRNAss, pre_miRNAs, and other non-coding RNAs were retrieved from Flybase database (as of Jan 2012, ftp://ftp.flybase.net/releases/FB2012_01/dmel_r5.43/fasta/). The exons from different alternative splicing isoforms were merged to find the maximum genome coverage regions per gene. An in-house perl script assigned each aligned short read to these merged transcription regions, and read counts per gene are the output. When a read is mapped to a region with more than one gene, i.e., one merged exon region overlapping with a non-coding gene, the count is split as equal possibilities into these two genes, half count for each. On average, only two percent of aligned reads are involved in gene overlapping regions.

Estimation of transcription level per gene per stage and differential expression analysis

Normalization term RPKM (Reads Per Kilo base pair per Million aligned reads) was first introduced to estimate the transcription level for short read data by normalizing transcript length and library size. In this study, we utilized the edgeR software package in *R* to find the normalization factors for libraries with various sizes (by the TMM (Trimmed Mean of M value) and upper quantile normalization methods). The edgeR method models short reads into negative binomial distribution and estimates the biological replicate variance (dispersion). Tag-wise dispersion estimation was performed in different groupings of libraries. At first, the 13 libraries came from 9 sample stages as niche, GB, S4, S8, S16, aly, can, EC16, LC. Among them, GB, S4, S8, and LC, each has two biological replicates. The two LC samples LC1 and LC2 appeared to reflect a gradual transition from early spermatocyte to late spermatocyte (shown in figure 2). Thus, in the "stage" grouping, samples from the same stage are grouped (GB1 with GB2, S4_1

with S4_2, and S8_1 with S8_2) except LC1 and LC2, generating 10 groups. The pseudo.alt and conc.group objects was calculated by edgeR using quantile normalization and maximum likelihood method. The pseudo.alt contains read counts after correcting the library size difference. We further normalized pseudo.alt with the length of the merged transcribed regions per gene. We introduced "corrected RPKM (cRPKM)" as pseudo.alt * 1e+09 / (length of merged transcripts)/ (common.lib.size) The common.lib.size was calculated from the calcNormFactors function of edgeR). The conc.group is a maximum likelihood estimate of the transcription level when two or more biological replicates are available. Since Conc.group has been normalized for common.lib.size, we only need to normalize for the merged transcript length. The normConc = conc.group *1e+09/(length of merged transcripts). For each gene, the "cRPKM" has one estimate for each of the 13 samples, while the "normConc" has 10 grouped estimates.

The differential expression genes are ranked when comparing a pair of group conditions, such as between niche to GB, in edgeR, with Benjamini (ref) multiple testing correction and adjusted P value <0.05.

Differential expression analysis

Due to single sample in several stages and highly correlated expression level, three transit amplifying stages (S4, S8, and S16) and early spermatocyte stages (EC16, aly, and can) were grouped for detecting differential expression with other groups (niche, GB, LC1 and LC2).

The Identification of Anonymously Transcribed Regions (ATRs)

There are approximately 5% of aligned reads outside the annotated transcribed regions (Flybase r543). And only approximately 3% of these un-annotated reads are located in introns, the rest (approximately 97%) are outside of the known gene regions. We utilized CUFFLINKS software to assemble these read to identify un-annotated transcribed regions (UTRs). We first merged all un-annotated reads from 13 samples and sifted for non-redundant ones. Then the

merged library was assembled by running CUFFLINKS with the –M switch masking the annotated region (the annotation file is from ftp://ftp.flybase.net/releases/FB2012_01/dmel_r5.43/gff/dmel-all-no-analysis-transcript-r5.43.gff). The assembled transcripts were further filtered for two criteria, first, the minimum transcript length is 100 nucleotides long; second, the adjusted RPKM (aRPKM) transcription level higher than 20.

Blastx search was performed following default parameter setting against NR protein sequence database. Only ATRs with at least 30 translated amino acids in the query-hit alignment and more than 30% sequence identity to another annotated protein sequence were reported as candidate coding transcripts. Blastn search followed default parameters to search ATRs in non-coding RNA (ncRNA) sequence library.

Other bioinformatics methods involved

The gene lists were compiled in a previous work from our lab including alternative splicing related genes, chromosome remodeling factors, histone modifying enzymes and transcription factors (list papers). MDS, hierarchical clustering, K-means clustering were performed by R software, with basis and heatmap library. Enrichment test of GO functional categories were performed utilizing online application from Flymine and DAVID.

Whole mount immunostaining

For each reaction, ~15 pairs of adult testis was dissected into 1xPBS containing tubes that were placed on ice. It is fixed in 4% formaldehyde/1x PBS solution for 30 minutes, followed by two 10 minutes washes in PBST (1xPBS/0.1% Triton X-100), leaving the third PBST washes for at least 30 minutes to aid permeabilization. Primary antibody diluted in 3%BSA/1xPBST solution was added and incubated at 4 °C overnight. Testes were washed three times in 1x PBST with 20

minutes incubation. Secondary antibody diluted in 3%BSA/1xPBST was added and incubated at 4 °C overnight. Testes were washed three times in 1x PBST with 20 minutes incubation. Testes were washed once in 1xPBS before mounted on slides in Vectashield mounting medium with DAPI (Vector laboratories, H1200). The primary antibodies used were: mouse anti-Armadillo(1:100; DSHB N2 7A1); rat anti-Vasa (1:40; DSHB from Spradling lab); rabbit anti-Zfh1 (1:4000; a gift from Ruth Lehmann, Skirball Institute of Biomolecular Medicine, NY, USA); and Chicken anti GFP (1:5000; abcam #13970) the testes were visualized using Zeiss Apotome and images were processed by Adobe Illustrator.

Results

Developing the Transcriptome Analysis using Single Germline Cyst (TASC) technique

To elucidate endogenous gene expression profiles in normally developing germ cells at discrete but continuous differentiation stages, we developed a novel method that we call Transcriptome Analysis using Single Germline Cyst (TASC). Germ cells at different developmental stages can be recognized by their distinct anatomical positions and morphological characteristics in *wild-type* testes (Fuller, 1998), which make it feasible to isolate single germ cell cysts at a particular stage of interest. Furthermore, we combined these advantageous features with cell type- and stage-specific GFP markers (Boyle et al., 2007; Chen and McKearin, 2003b; Chen et al., 2005; Santel et al., 1997) (Fig. 2-1A and Materials & Methods). Specifically, hub-specific GFP marker [*unpaired* (*upd*)-*Gal4>UAS-GFP* (Boyle et al., 2007)] was used to isolate Niche sample, spermatogonia-specific Bag-of-marbles (Bam)-GFP marker (Chen and McKearin, 2003b) was used to isolate different spermatogonial cysts, spermatocyte-specific Spermatocyte arrest (Sa)-GFP marker (Chen et al., 2005) was used to isolate the spermatocytes samples, and spermatid-specific Don juan (Dj)-GFP marker was used to isolate post-meiotic spermatid samples. Together, 10 distinct stage single germline cyst samples were physically isolated from *wild-type* testes: Niche (N), Gonialblast (GB), 2-cell Spermatogonia (S2), 4-cell Spermatogonia (S4), 8-cell Spermatogonia (S8), 16-cell Spermatogonia (S16), Early Spermatocyte (EC), Large Spermatocytes (LC), Round Spermatid (RS), and Elongating Spermatid (ES). In addition, two mutant spermatocyte samples were isolated from *cannonball* (*can*) (Hiller et al., 2001; Lin et al., 1996) and *always early* (*aly*) (White-Cooper et al., 2000) mutant testes, which represent the earliest spermatocyte stage due to genetic arrest. We then systematically profiled transcriptome at each stage of spermatogenesis by performing RNA-seq using isolated germline cysts (Fig. 2-1B), using an adapted amplification method for single cell transcriptome analysis (Kurimoto et al.,

2007; Tang et al., 2009). The purity of the sample offered an unprecedented opportunity to study transcriptional dynamics at both the genome-wide level and individual gene resolution.

The final reads (28 million in total, from 13 sample libraries) were assigned to annotated genes based on their merged transcript regions (Gan et al., 2010a). The original RPKM (Gan et al., 2010a) normalized total read counts across libraries and transcripts' length. Here we combined RPKM computation with TMM and upper-quantile normalization methods to better control variation among different libraries. We then used post-normalization read counts (pseudo.alt generated by EdgeR) followed by normalization with merged transcript length to get corrected RPKM (cRPKM) values (see Materials and Methods).

Overview of the TASC data: dynamic transcriptional changes during spermatogenesis

We first set up a threshold for a gene to be called "actively transcribed". We compared cRPKM value with published microarray data (Chen et al., 2011; Terry et al., 2006) (Figure 2-2), which showed that genes with cRPKM greater than 30 contain most "present"-call genes in microarray data. Therefore we used cRPKM=30 as a cutoff line for expressed genes, although this probably represents a very stringent cutoff because the sensitivity of RNA-seq technique is higher than microarray and can detect lowly expressed genes (Wang et al., 2009). Using this standard, 8,551 out of 18,000 annotated *Drosophila* genes were found to be transcribed at ≥ 1 stage during *Drosophila* spermatogenesis (Fig. 2-1C). We then examined the overall dynamics of these 8,551 genes along cellular differentiation program, by performing clustering analysis. Among the ten clusters identified using K-means algorithm, only one cluster (413 genes) showed overall downregulation (Fig. 2-3A) while five clusters (3,972 genes) showed overall upregulation (Fig. 2-3B-F). The remaining four clusters (4,539 genes) showed overall stable level of gene expression (Fig. 2-3G-J). Based on these analyses, we concluded that gene activation is likely the major mode of differential gene expression during spermatogenesis. On the other hand, stable

level of transcript may also be due to post-transcriptional regulation, which has been proposed as a unique molecular mechanism in regulating germ cell function (Cinalli et al., 2008; Rangan et al., 2008; Seydoux and Braun, 2006).

To gain a global picture of the transcriptome change during spermatogenesis, we performed unsupervised clustering using spearman's correlation coefficient generated between each pair of the 13 samples (Fig. 2-1D). We identified three discrete clusters representing three major transcriptional transitions, which coincide with morphological changes at the cellular level (Fuller, 1998). The first transition was from Niche to GB samples, the second one occurred from the transit-amplifying S16 to early spermatocyte sample (i.e. *aly*, *can*, and EC samples), and the third one was from early to late spermatocyte samples. Based on this analysis, it is obvious that the two LC samples picked based on their morphology were not identical. The LC1 sample had a molecular signature closer to early spermatocyte than LC2, therefore should represent an intermediate stage between early and late spermatocytes. Because spermatocyte takes a long time to transition from early to late stage (3-4 days), morphological criterion cannot pinpoint the exact maturation stage. On the other hand, this intermediate stage provides a higher resolution dataset to understand spermatocyte maturation, an elongated G2 phase of meiosis I in preparation for meiotic divisions and terminal differentiation. These three transcriptional transitions and their biological relevance will be discussed in more details.

In addition to these three major transcriptional transitions, dynamic gene expression was also detected between each two consecutive differentiation stages. Because S2 to S16 spermatogonial stages share a similar transcriptome, as well as the three early spermatocyte samples including *aly*, *can* and EC, we combined them into two separate groups. Based on the number of differentially expressed genes, the transition from spermatogonia to early spermatocyte represented the most dramatic one with 503 downregulated genes and 3,009 upregulated genes (Fig. 2-1E). Our finding is consistent with previous studies comparing testes enriched with

spermatogonial cells to *wild-type* testes containing spermatocytes (Chen et al., 2011; Gan et al., 2010a; Terry et al., 2006).

The first transcriptome transition between niche and GB samples

In *Drosophila* testis, GSC divides asymmetrically most of the time and the self-renewed GSC remains physically attached to the hub cells *via* adherens junctions, whereas the other daughter cell GB is displaced away from the hub and initiates transit-amplification (Yamashita et al., 2003). However, beside the anatomical difference between GSC and GB, it is unclear whether they have different molecular characteristics. Our recent study revealed that the chromatin state between GSC and GB is different (Tran et al., 2012), suggesting that they may have distinct gene expression profiles. Here we sought to gain insight into this by comparing the TASC data between Niche and GB samples. One caveat in this comparison is the heterogeneity of the Niche sample, which contains hub cells and CySCs in addition to GSCs. However, we reasoned that the upregulated genes in the GB sample should represent *de novo* transcribed genes in GB or accompanying cyst cells; whereas enriched transcripts in the Niche sample would shed more light on important regulators required for stem cell or niche functionality. Both lines of information will help understanding how stem cells maintain their unique property and how their immediate daughter cells initiate the differentiation program.

We first plot the mean transcription level of genes in Niche vs. GB samples [Fig. 2-4A, differential expression with statistical significance (P-value <0.05, based on negative binomial model with Benjamini multiple testing correction) in red; blue lines represent a 10-fold difference]. In total, we identified 381 niche-enriched genes (10-fold higher in the niche sample), among which 91 showed statistical significance (P< 0.05). Ontology assay on niche-enriched genes revealed proteases (33 genes, Fig. 2-4B) as the top category with biological significance. Among these proteases we found Spatzle-Processing Enzyme (SPE). SPE is a serine-type

endopeptidase, which is known to process and activate Spatzle (Spz), the ligand of the Toll signaling pathway (Mulinari et al., 2006). We found that knocking down SPE (*UAS-SPE dsRNA*) in germ cells using the *nanos* (*nos*)-*Gal4* (Van Doren et al., 1998) driver caused a dramatic germ cell loss phenotype in testis (Fig. 2-4C). Following germ cell loss, Zfh-1-positive early cyst cells including CySCs became overpopulated in the niche. By contrast, when the same *UAS-SPE dsRNA* transgene was driven by a hub-specific *upd-Gal4* driver (Boyle et al., 2007; Leatherman and Dinardo, 2010) or a cyst cell-specific *c587-Gal4* driver (Manseau et al., 1997), no obvious phenotype could be detected at the niche (Fig. 2-4D-E). Based on these cell type-specific knockdown results, we concluded that SPE is required in early germ cells to maintain GSCs in the niche. Therefore our Niche-enriched gene dataset can be used as an initial molecule screen to identify intrinsic factors required for GSC maintenance. Further studies will be warranted to elucidate the biological significance of enriched Toll signaling pathway components in the testis niche.

We also identified differentially expressed genes that are more enriched in the GB sample compared to the niche samples. Specifically, we identified 742 GB-enriched genes (10-fold higher in the GB sample), among which 201 showed statistical significance ($P< 0.05$). When we performed gene ontology assay, we found that chromatin factor category is the top one with biological significance (Fig. 2-4F). We hypothesize that these chromatin factors may reset chromatin landscape to allow transcription of differentiation genes. Interestingly, once these chromatin factors are turned on, for most of them the expression level is maintained stable throughout the rest stages. It is possible that changes at the chromatin level precede and prepare cells for turning on differential gene expression program for cell fate commitment.

Transit-amplifying spermatogonial cells share similar transcriptome

In most adult stem cells lineages, a small number of stem cells provide a large pool of differentiated cells required for tissue homeostasis. Therefore proliferation of progenitor cells, the immediate descendent of a stem cell, is key to normal stem cell function. In *Drosophila melanogaster* testis, GB undergoes exactly four round of mitosis to expand from one cell to 16 cells, a stage called transit-amplification (Fuller, 1998). Using the combination of Bam-GFP marker (Chen and McKearin, 2003b) and cyst size, we isolated each transit-amplification stage from S2 to S16 and profiled their transcriptomes. Pair-wise comparison showed >0.9 correlation coefficient between any two samples, including GB and spermatogonial samples (Fig. 2-5A). Furthermore, all transit-amplification samples share a common pool of expressed genes (2,990 genes, cRPKM≥ 30, Fig. 2-5B). Ontology assay on these commonly expressed spermatogonial genes showed biological categories such as ribonucleoprotein biogenesis and DNA replication (Fig. 2-5C), which are required for cells to undergo rapid proliferation. Similar transcriptome among transit-amplifying cells (GB, S2, S4, S8 and S16) may provide a molecular basis to explain their similar behavior during both de-differentiation and differentiation processes. Dedifferentiation of spermatogonial cells has been identified as a mechanism to replenish lost stem cells during genetic ablation or aging (Brawley and Matunis, 2004; Cheng et al., 2008; Sheng et al., 2009; Sheng and Matunis, 2009; Wong and Jones, 2012). On the other hand, transition from spermatogonia to spermatocyte pre-maturely or at a later point does not affect proper differentiation, indicating their equal differentiation potential (Eun et al., 2013; Insco et al., 2012; Insco et al., 2009; Parrott et al., 2011).

Although the overall transciptome among transit-amplification stage samples are very similar, we also found a few stage-specific genes (Fig. 2-5D). Ontology analysis did not show any category with specific biological function, which is probably due to the very small number of

genes. However, these genes could be used as cellular markers to precisely label each of the transit-amplification stage.

The second transcriptome transition between spermatogonial and spermatocyte samples

Up to date, the best understood transcriptional switch during *Drosophila* spermatogenesis is from the 16-cell spermatogonia to spermatocyte, representing the change from mitotic to meiotic program (Lim et al., 2012). Previous studies on key regulators, such as the testis-specific homologs of TATA-binding protein-associated factors (tTAFs) and the testis-specific meiotic arrest complex (tMAC), demonstrate an orchestrated mode of transcriptional activation in spermatocytes (Ayyar et al., 2003; Beall et al., 2007; Chen et al., 2011; Hiller et al., 2004; Hiller et al., 2001; Jiang et al., 2007; Jiang and White-Cooper, 2003; Lin et al., 1996; Perezgasga et al., 2004; White-Cooper et al., 2000; White-Cooper et al., 1998). In addition, spermatocyte-specific tTAF and tMAC proteins antagonize Polycomb group (PcG) transcriptional repressive complex to derepress differentiation genes that are silenced in spermatogonial cells (Chen et al., 2005; Chen et al., 2011; Gan et al., 2010b). Based on these results, it has been hypothesized that there are at least two transcriptional transitions in spermatocytes: one at the spermatogonia-to-spermatocyte transition, which turns on transcriptional and chromatin regulators to set up the chromatin landscape, in preparation for the next one, from early to late spermatocyte, when terminal differentiation genes are turned on (Chen et al., 2011). It is very likely that genes turned on at the second transition are target genes controlled by genes turned on at the first transition. We first checked the tTAF and tMAC expression pattern in our dataset. We confirmed that most of their transcription is low in mitotic program but is abruptly upregulated in early stage spermatocyte (Fig. 2-6A). We also confirmed that their transcription is independent of their own products, because of the highly detectable transcript in either *can* or *aly* mutant spermatocyte samples (Fig. 2-6A). These are consistent with previous reports using Northern blot, *in situ*

hybridization (Hiller et al., 2001; White-Cooper et al., 1998), microarray (Chen et al., 2011) and RNA-seq (Gan et al., 2010a) analysis using dissected testes.

Next, we asked how the expression pattern of tTAF and tMAC genes applies to other key transcriptional regulators, such as transcription factors, alternative splicing factors, histone modifying enzymes, and chromatin remodeling factors. Previous results using *bam* or *benign gonial cell neoplasm* (*bgcn*) mutant testes enriched with mitotic spermatogonial cells to *wild-type* testes (Chen et al., 2011; Gan et al., 2010a; Terry et al., 2006) revealed enriched alternative splicing factors, histone modifying enzymes and chromatin remodeling factors in mitotic cells. Interestingly, analysis of annotated alternative splicing factors, histone modifying enzymes and chromatin remodeling factors revealed a bi-modal pattern in which either a low-to-high or high-to-low switch is observed at the S16 to early spermatocyte transition (Fig. 2-6C). By contrast, we found more annotated transcription factors showed upregulation than downregulation in the switch from S16 to early spermatocyte stages, consistent with the tTAF and tMAC expression pattern and also with the major activation mode of differential gene expression during spermatogenesis (Fig. 2-3). The few transcription factors (20%, 30/154) that are downregulated from S16 to early spermatocyte stage could be responsible for shutting down mitosis-specific genes to avoid over-proliferation and tumorigenesis.

This mitosis-to-meiosis switch needs to be tightly regulated. Any misregulation of this transition may lead to infertility due to insufficient germ cells or testicular tumor due to spermatogonial overproliferation. Although *bam* and *bgcn* genes are key regulators for this transition (Gonczy et al., 1997; Insco et al., 2012; Insco et al., 2009; McKearin and Spradling, 1990), our genome-wide and single-cyst transcriptome analysis have revealed many more key transcriptional regulators that showed a dynamic bi-modal pattern at the mitosis-to-meiosis switch. Functional analysis of these candidate genes will shed more light on the molecular mechanisms underlying exit of mitosis and onset of meiosis. To our current knowledge, this

decision makes an irreversible commitment on lineage specification because spermatocytes lose dedifferentiating potential (Sheng et al., 2009).

The third transcriptome transition between early and late spermatocyte samples

The maturation of spermatocytes takes the longest time during the entire spermatogenesis, which lasts for 3-4 days and involves 25-fold growth in volume. Our dataset provides a high resolution view of the transcriptome change during spermatocyte maturation with *aly*, *can* and EC samples representing early spermatocyte, LC1 representing an intermediate while LC2 representing the very late and mature spermatocyte stage (Fig. 2-1E). Furthermore, our clustering analysis confirmed that *aly* mutant spermatocytes are arrested at an even earlier time point than *can* mutant spermatocytes which is consistent with prediction based on previous reports (Ayyar et al., 2003; Beall et al., 2007; Chen et al., 2011; Hiller et al., 2004; Hiller et al., 2001; Jiang et al., 2007; Jiang and White-Cooper, 2003; Lin et al., 1996; Perezgasga et al., 2004; White-Cooper et al., 2000; White-Cooper et al., 1998). Therefore our TASC analysis provided a progression of spermatocyte maturation in the following order: *aly* mutant spermatocyte→ *can* mutant spermatocyte→ *wild-type* EC spermatocyte→ *wild-type* LC1 spermatocyte→ *wild-type* LC2 spermatocyte (Fig. 2-6B).

Using this linear progression model, we proceeded to analyze differential gene expression at every single spermatocyte sample and found a sequential activation mode. We hypothesize that genes expressed at early spermatocyte stage encode upstream regulators which subsequently turned on genes required for meiotic divisions and terminal differentiation.

Lack of dosage compensation in germline cysts throughout *Drosophila* spermatogenesis

Dosage compensation equilibrates X chromosomal and autosomal gene expression between male and female by hyperactivation of X-chromosomal gene transcription (Deng et al.,

2011; Gupta et al., 2006; Parisi et al., 2003). Hyperactivation of X-linked genes requires dosage compensation complex (DCC) in male somatic cells, which generates H4K16ac active histone modification and promotes RNA polymerase II elongation (Gelbart et al., 2009). Our previous RNA-seq data using gonadectomized flies (adult flies with testis or ovary removed, labeled as "carM" and "carF" in Fig. 2-7A) showed little difference between X-linked genes and autosomal genes in their median expression level, confirming that dosage compensation occurs in somatic cells.

Despite extensive knowledge about dosage compensation in somatic cells, whether and how X-chromosomal genes are compensated in male germline is not fully understood. Analysis fof microarray and RNAseq data using isolated *wild-type* testes showed significant lower expression level of X-linked genes compared to autosomal genes (Fig. 2-7A), suggesting lack of dosage compensation. By contrast, analysis of *bam* mutant testes showed comparable level between X-linked genes and autosomal genes, suggesting existence of dosage compensation. Because *wild-type* testes are enriched with meiotic germ cells whereas *bam* testes mainly contain mitotic germ cells, these data suggest dosage compensation may exist in mitotic germ cells in *Drosophila* testes. However, using the entire testes unavoidably include other cell types in addition to germ cells, which may obscure the results.

Here, using our TASC RNA-seq data, we revisited dosage compensation assay. Even though our isolated cyst samples still contained somatic cells, most of them are enriched with germ cells, for example, in S16 and all spermatocyte samples, the cyst cell to germ cell ratio is 2 to 16. In addition, because dosage compensation exists in somatic cells, if we find strong evidence suggesting low or no dosage compensation in our samples, it should be due to lack of dosage compensation in germ cells. In order to prevent any bias introduced by the unique and dynamic gene expression program during spermatogenesis, we carefully restricted our analysis to non-differentially-expressed, non-testis-specific and non-male-specific genes (see Method). We

showed that in all spermatogonial stages ("GB", "S4", "S8", "S16" in Fig. 2-7A), X-linked genes are expressed at a significantly lower level than autosomal genes, different from the previous results obtained using *bam* testes sample. There are at least three possibilities leading to this discrepancy: first, previous data analysis did not exclude those sex-biased or stage-biased genes, which may change the results significantly. Second, *bam* testes are heterogeneous and contain many other somatic cell types such as cyst cells, muscle sheathe cells, and pigment cells. Third, *bam* mutant germ cells are different from *wild-type* spermatogonia cells, which is shown by the distinct transcriptome of *bam* testes compared to all spermatogonial cyst samples (Fig.2-6D). We further confirmed that none of the meiotic cyst sample showed evidence of dosage compensation ("aly", "can", "EC", "LC1", "LC2" in Fig. 5A). Finally, the niche sample showed highly compensated X-linked genes, probably due to presence of somatic hub cells and CySCs in this sample ("niche" in Fig. 2-7A). We summarized dosage compensation index defined by (2- A:X) [0 indicates no compensation; 1 indicates 2-fold upregulation of X-linked gene expression] of our TASC samples and previous tissue samples (Fig. 2-7B). Although most of the previous tissue samples showed compensation, only the niche sample showed compensation among our TASC samples, suggesting male germ cells are lack of dosage compensation when pure starting material with much less contribution of somatic cells is used.

New features of *Drosophila* germline transcriptome

When we aligned all short reads uniquely to *Drosophila* genome (Flybase r5.43), approximately 6-7% of total reads were located outside annotated transcriptome including both mRNAs and non-coding RNAs. We assembled these reads into newly transcribed regions (NTRs), using the Cufflinks program (Trapnell et al., 2010). We identified a total of 1,626 NTRs, which have active expression and considerable transcript length (see Methods). We further

excluded previously identified NTRs using RNA-seq of tissues, including 454 NTRs in somatic tissue and 328 NTRs in testes (Gan et al., 2010a). The remaining 844 NTRs constitute a set of testis-specific transcripts that have not been documented in the current Flybase database (r543).

The transcriptional profile of these 844 NTRs during male germ cell differentiation is illustrated in Figure 2-8A. Approximately 75% of these NTRs show stage-specific expression pattern (Fig. 2-8A). Using hierarchical clustering, five major gene expression clusters were found for these NTRs, which three of the clusters show abrupt changes at the spermatogonia-to-spermatocyte transition. This pattern is consistent with other annotated genes in our dataset (Fig. 2). The differential expression pattern of testis-specific NTRs suggests that they are relevant to the mitosis-to-meiosis transition in male germline lineage. We next analyzed the chromosomal distribution of NTRs and found a significant enrichment on X chromosome, 4th chromosome, 2Rhet, 3Lhet and Yhet heterochromatic regions, as well as U regions (unassembled scaffolds) (Fig. 2-8B). We also found that the majority of the NTRs are near annotated gene region (919/1,172 as "near-gene" group, with 909 NTRs within 1kb flanking sequence and 10 NTRs in the intronic region of annotated genes). Only 253 out of 1,172 testis-specific NTRs are located at the gene desert region ("gene-desert" group: no overlap with any annotated gene with 1kb extensions at both 5' and 3' ends). We further validated that these NTRs are indeed transcribed by performing RT-PCR experiments (21/22 "gene-desert" NTRs validated, Fig. 2-8C).

To understand the potential biological functions of NTRs, we BLAST their sequences in the translated non-redundant protein database (nr) and non-coding RNA (ncRNA) sequence library of Drosophila (Flybase r5.43). Higher percentage of "gene-desert" NTRs (31.2%, 79 out of 253) had homology to translated protein sequences than "near-gene" NTRs (6.0%, 55 out of 919), suggesting they are more likely protein-coding (Fig. 2-8D, P = 2.8e-15 by Fisher Exact Test). We next analyzed the length distribution of NTRs (Fig. 2-8E): "near-gene" NTRs tend to have shorter length at 100-200 nt, typically for pre-miRNAs, snRNAs or snoRNAs. It is also

possible that these “near-gene” short NTRs are exons due to alternative 5’ transcription start site or alternative 3’ transcription end site usages. We screened for junction reads covering both the “near-gene” NTRs and a nearby annotated transcript. Indeed we found that 10 putative NTRs are possibly unannotated exons: in two cases, the newly identified exons will extend transcripts 4913 nt and 13,821 nt, respectively. On the other hand, “gene-desert” NTRs showed two groups with different transcript length at 100-200 nt and 900 nt, respectively. When we analyzed the longer “gene-desert” NTRs, only 89 have homology with protein coding sequences, the rest could be non-coding RNAs including 33 candidate long non-coding RNA (lincRNAs) (>2000 nt RNA with no known protein coding potential).

Discussion

In this paper we used a new isolation method to obtain pure germ cells recognized by their precise differentiation stages. The usage of tissue from wild-type animal avoids potential complications in interpreting data obtained from tissues obtained from mutant animals. We combined these pure samples with RNA-seq technology, which has lowered the threshold of starting materials required for mRNA sequencing to nanogram scale. To our knowledge, this dataset represents the first entire lineage transcriptome from an endogenous adult stem cell system. This high resolution transcriptome database has provided lots of biological insights to understand Drosophila male GSC differentiation, which will be advance the knowledge in germ cell biology and stem cell biology fields. Nest we will discuss our findings and their implications.

The GB is committed for differentiation

One surprising finding from our TASC dataset is that the GB sample has a distinct transciptome compared to the Niche sample, even though GB and GSC are the immediate two daughter cells from one asymmetric cell division. It is also interesting that many chromatin regulators have increased expression in the GB sample, suggesting it may have a very dynamic chromatin landscape. Consistent with this, recently we found that GB undergoes a rapid histone 3 (H3) turnover to replace preexisting H3 with newly synthesized H3, when the two cells (i.e. GSC and GB) are still connected by the spectrosome structure and undergo DNA replication synchronously (Tran et al., 2012). We hypothesize that such a dynamic histone turnover could lead to the dramatically changed transcriptome in GB, although at this moment we do not know whether increased expression of chromatin regulators is the causal reason or a consequence of the histone replacement. Nevertheless, this step could be crucial to reset the entire chromatin in GB to prepare it for differentiation. Our TASC data also revealed that from GB to 16S, the

transcriptome remains stable, suggesting that transit-amplification acts to simply expand progenitor cell number without dynamic change at the chromatin or transcriptional levels.

These findings will explain several observations and raise interesting discussion as well. First, it will explain why decisions made as early as GSC asymmetric division can affect gene expression in meiotic spermatocytes. For example, a recent study showed that sex chromosomes have biased segregation patterns during asymmetric cell division of male GSC (Yadlapalli and Yamashita, 2013). Interestingly, the authors found that in the mutants where the bias was randomized, ectopic gene expression could be observed in spermatocytes. Second, it will also resolve findings from several labs demonstrating that transit-amplifying cells can transit to spermatocytes without any obvious differentiation defect (Eun et al., 2013; Insco et al., 2012; Insco et al., 2009; Parrott et al., 2011). However, it will also raise a question whether dedifferentiated spermatogonial cells can function as *bona fide* GSCs. For example, previous studies showed that dedifferentiated GSCs tend to have misoriented centrosomes (Cheng et al., 2008) and lose biased sex chromosomal segregation (Yadlapalli and Yamashita, 2013). It will be interesting to find out whether dedifferentiated GSCs can restore the transcriptional profile, a hallmark for successful reprogramming.

Technical discussions

Since germ cell cysts consist of synchronously dividing and differentiating germ cells plus two encapsulating cyst cells, we tried to further lyse the two enveloping cyst cells in order to get germ cells only. But such a treatment led to collapse of the cyst and failed to give us sufficient amount of mRNA to construct libraries. We also tried to obtain hub cell only sample (i.e. Niche sample without GSCs or CySCs), but such a sample also gave out very low reads when sequenced. There are two possible reasons leading to this: first, the lysis treatment could be too detrimental to retain intact transcripts; or, it could be that post-mitotic hub cells have overall low

transcriptional activity because they are mainly known to generate signaling pathway ligands and cell-cell adhesion molecules. In fact, although post-meiotic transcripts were reported (Barreau et al., 2008a; Barreau et al., 2008b), there are very few genes identified and it is unclear the level of transcripts present. And in general, the two post-meiotic samples we have, the RS and ES samples, showed very few RNA-seq reads, probably due to the general shut-downed transcription in post-meiotic germ cells.

Technology advancement has provided us with unprecedented opportunity to understand some critical questions. For example, combining our dataset with single cell ribosomal profiling or proteomics studies when those techniques become available, will further reveal how post-transcriptional mechanisms regulate germ cell differentiation at every single stage during spermatogenesis.

Figures and Tables

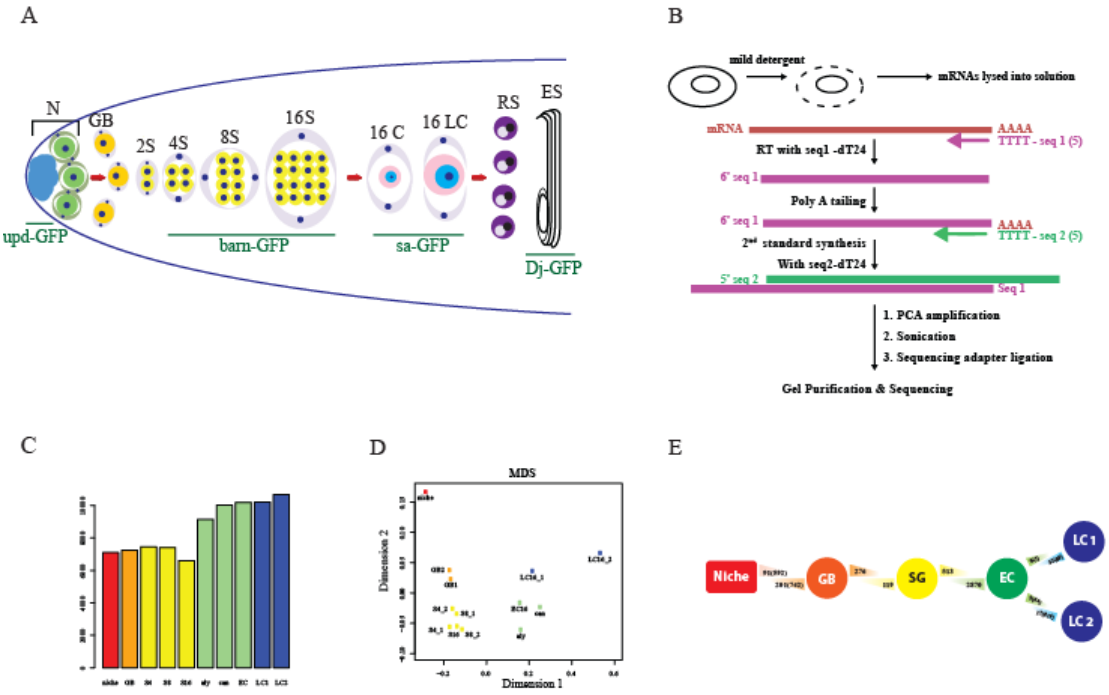


Figure 2-1: Global overview of the transcriptome library. A. Cell and stage specific GFP markers and morphology. **B.** RNAseq workflow to generate library from isolated sample. **C.** Number of genes expressed in consecutive stages. **D.** Clustering of spearman correlation coefficient for each library. **E.** Dynamic changes in gene expression between consecutive stages.

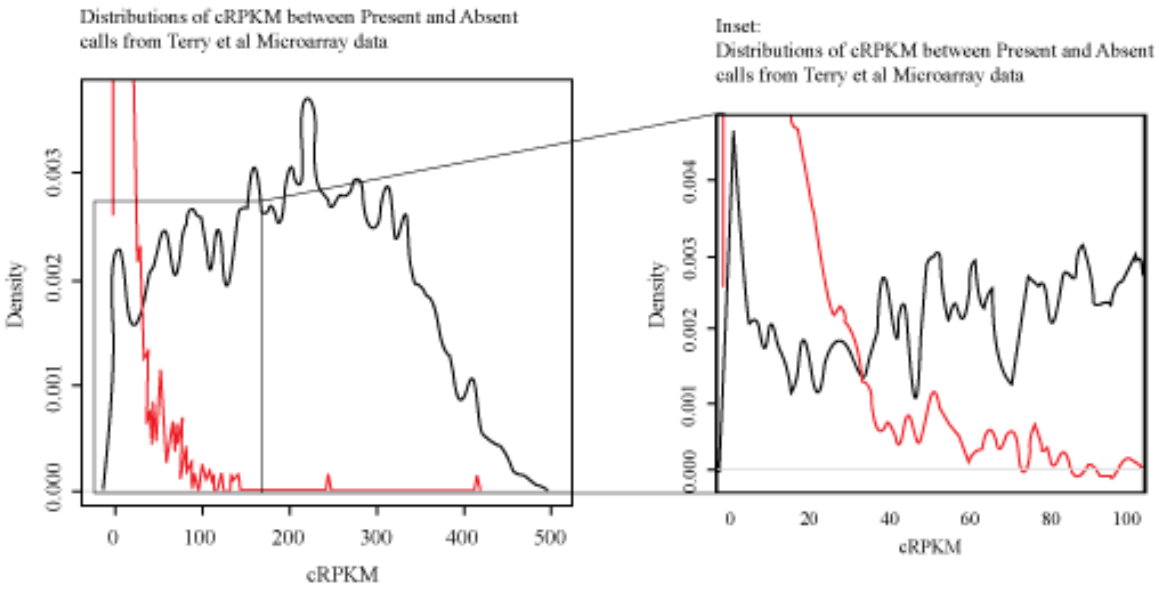


Figure 2-2: Distribution of present and absent call from published microarray data. The absent call from microarray data are plotted as Red and the present data are plotted as Black. Inset shows that the present and absent lines intersect at cRPKM=30.

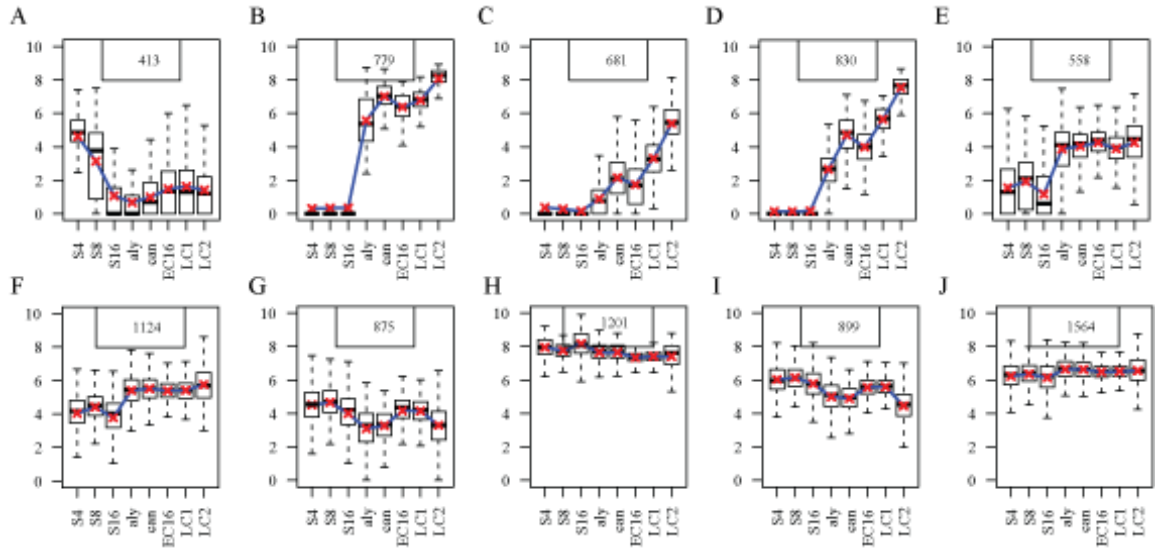


Figure 2-3: Clustering analysis using K-means algorithm on all transcribed genes across spermatogenesis stages.

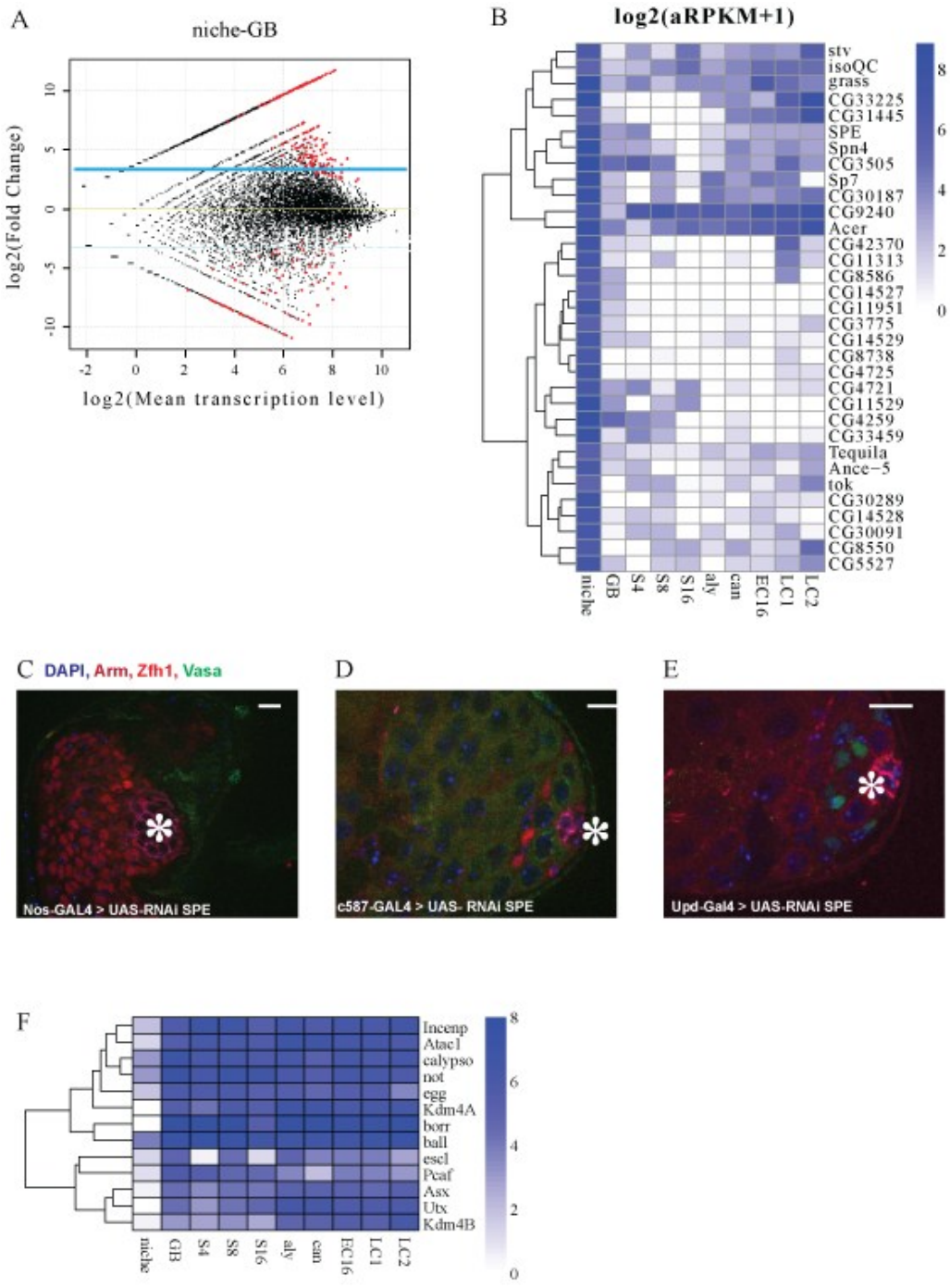


Figure 2-4: Transcriptome transition between niche and GB samples A. Mean transcription level of expressed genes in niche and GB. Blue line represent 10 fold change and red dots overlay is statistically significant differentially expressed genes. **B.** Ontology assay on Niche-enriched genes identified Protease as biologically significant category. **C.** Knockdown of SPE using

germline specific (nos-gal4) showed germ cell loss. Germ cells are labeled using Vasa (green); early somatic cells are labeled using zfh1 (red); hub cells are labeled using Armadillo (magenta); and DAPI labeled DNA (blue). **D.** Knockdown of SPE using early cyst specific driver (c587-gal4) show normal niche morphology. **E.** Knockdown using hub specific driver (Upd-gal4) showed normal niche morphology. **F.** Ontology assay on Gonialblast-enriched genes identified chromatin remodelers and histone modifying enzyme as biologically significant categories.

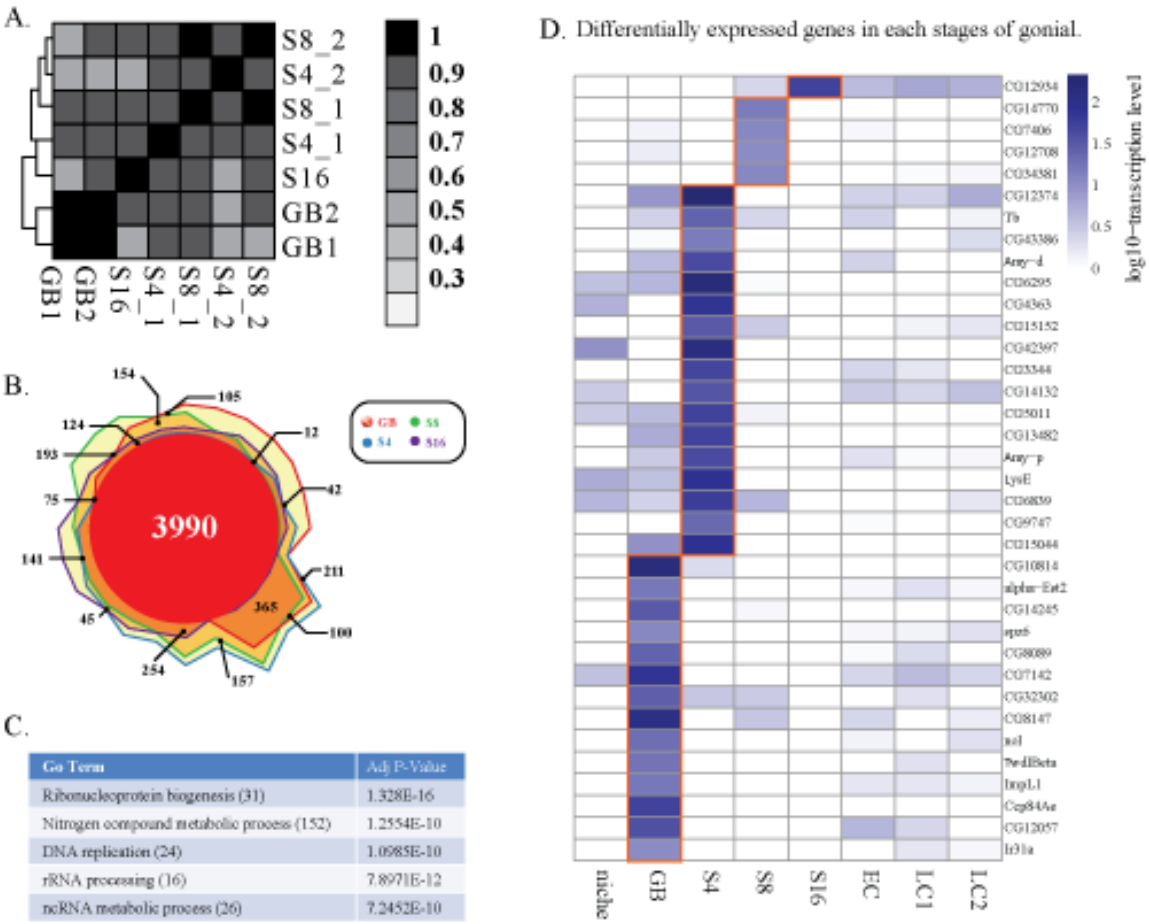


Figure 2-5: Transit amplifying spermatogonial cells share similar transcriptome. A. Heat map showing pair-wise correlation between spermatogonia samples **B.** Diagram showing the overlap between genes in each spermatogonia samples **C.** Ontology analysis for genes shared by

all spermatogonia samples. **D.** Heat map of differentially expressed genes specific to each spermatogonia samples

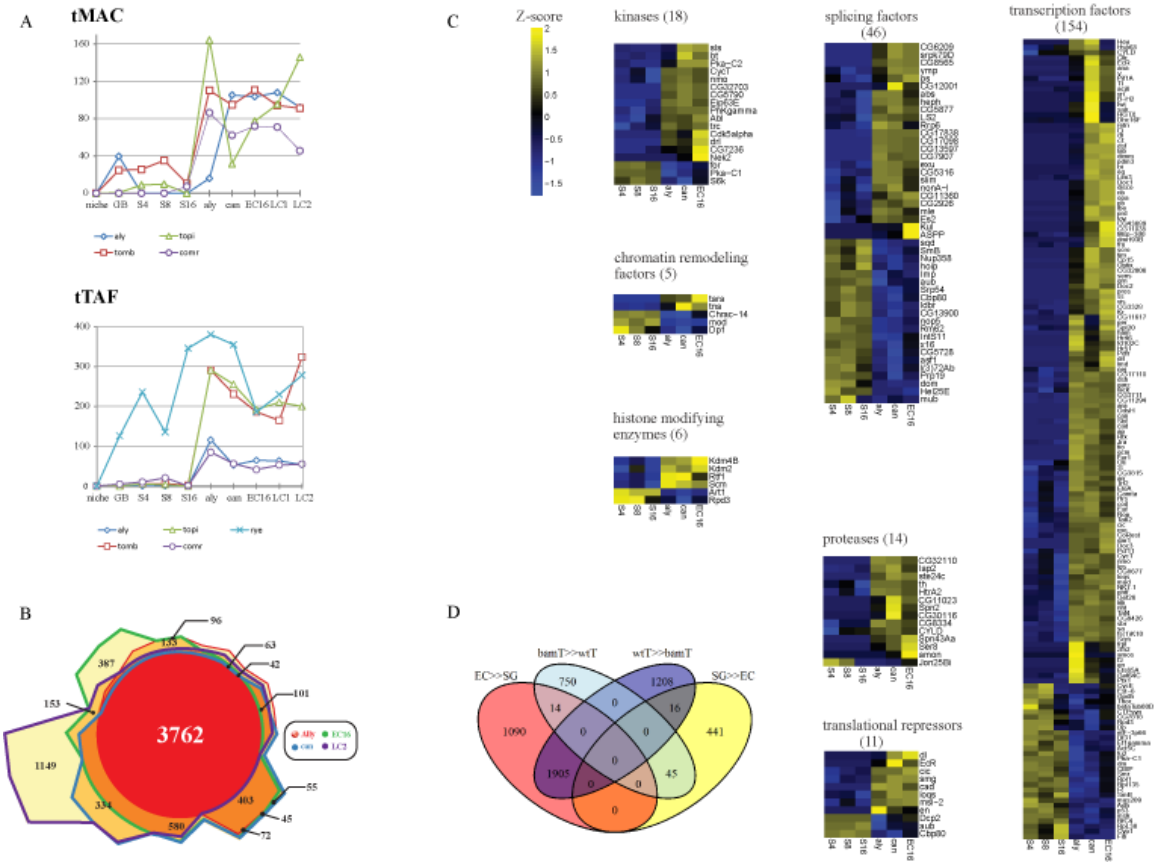


Figure 2-6: Transcriptome transition between spermatogonial and spermatocyte samples. **A.** Expression pattern of tMAC and tTAF component. **B.** Venn diagram showing the sequential activation of spermatocyte maturation **C.** Heat map showing bimodal expression pattern at spermatogonial to spermatocyte switch. **D.** Venn diagram comparing all spermatocyte samples, bam, and wildtype tissue samples.

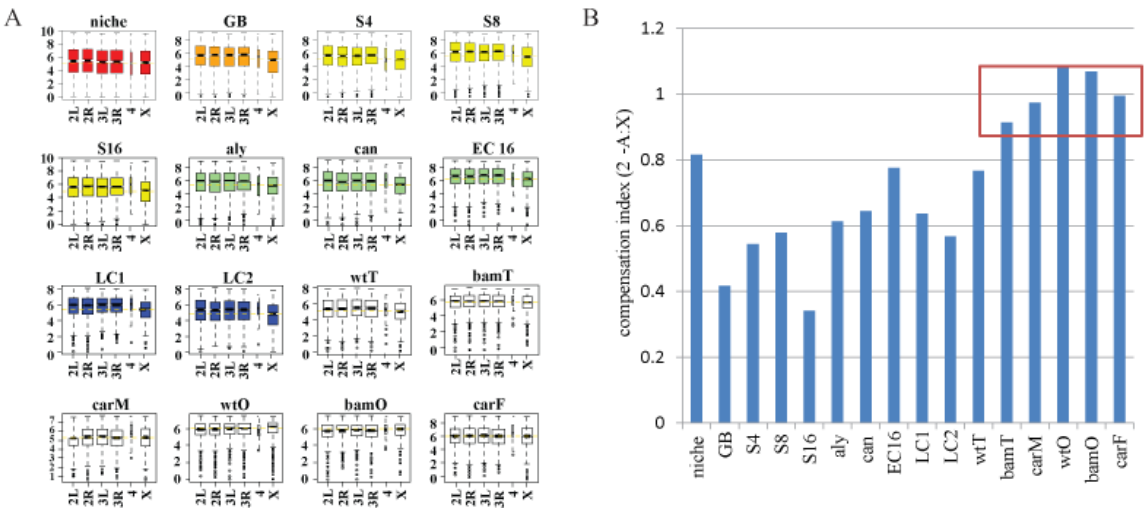


Figure 2-7 : Lack of dosage compendation in germline cysts throughout *Drosophila* spermatogenesis. A. Expression level of genes located on each of the chromosomes. **B.** Compensation index for datasets. Red box showed samples with significant compensation.

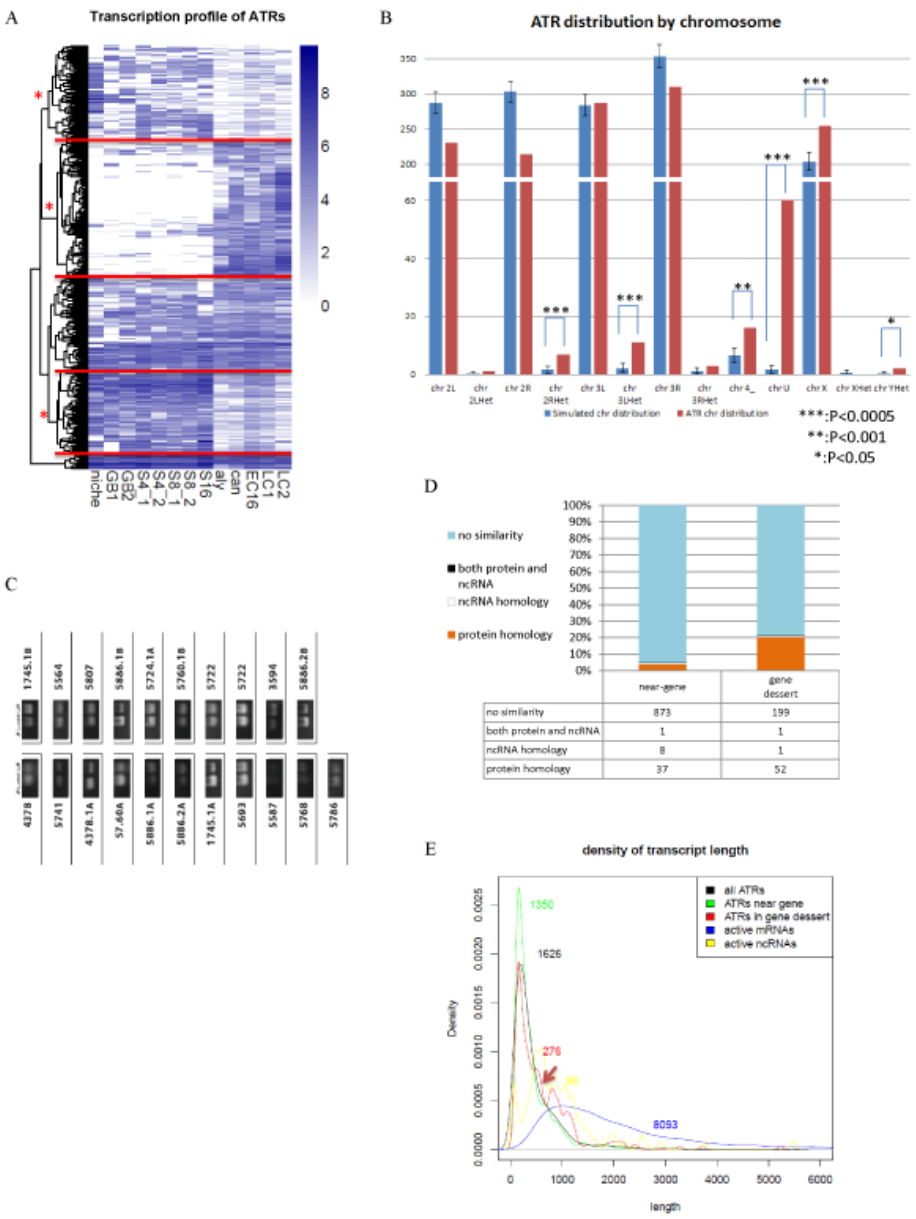


Figure 2-8: New Features of *Drosophila* germline transcriptome. A. Heat map shows that NTRs are expressed stage specifically. **B.** NTR distribution by chromosome **C.** PCR validation

show the expression of NTR in testis sample. **D.** BLAST result for NTRs. **E.** Length distribution of NTRs

sample	total read sequenced (a)	uniquely mapped (b=c+d)	% (b/a)	mapped onto genome sequence (c)	exon junction region (d)	% (d/b)	non-redundant uniquely mapped (e)	% (e/b)	in annotated transcript exons (f)	% (f/e)	mapped outside annotated exon region (g)	% (g/e)	mapped onto intron region	mapped onto intergenic region
niche	10250514	5902780	58%	5843891	58889	1.0%	1284576	22%	1209317	94%	75259	5.9%	1110	74149
GB1	8727415	4540540	52%	4481181	59359	1.3%	1667631	37%	1606184	96%	61447	3.7%	285	61162
GB2	7561213	4857912	64%	4787363	70549	1.5%	1517874	31%	1447346	95%	70528	4.6%	839	69689
S4_1	7333933	3504177	48%	3453952	50225	1.4%	1161121	33%	1102010	95%	59111	5.1%	662	58449
S4_2	10627923	7554559	71%	7505812	48747	0.6%	2137743	28%	2054508	96%	83235	3.9%	747	82488
S8_1	14755082	7225177	49%	7169631	55546	0.8%	1845556	26%	1769253	96%	76303	4.1%	650	75653
S8_2	15035629	11659858	78%	11633597	26261	0.2%	2902198	25%	2771649	96%	130549	4.5%	1274	129275
S16	8408805	2408098	29%	2393346	14752	0.6%	857356	36%	807763	94%	49593	5.8%	484	49109
aly	4147016	3430729	83%	3392315	38414	1.1%	1788880	52%	1709236	96%	79644	4.5%	917	78727
can	5556415	4478958	81%	4438218	40740	0.9%	2295652	51%	2170046	95%	125606	5.5%	2045	123561
EC16	25460503	17844021	70%	17785436	58585	0.3%	3892719	22%	3597589	92%	295130	7.6%	4804	290326
LC1	12544360	10432885	83%	10412065	20820	0.2%	3381650	32%	3209121	95%	172529	5.1%	2122	170407
LC2	10738832	8871846	83%	8848024	23822	0.3%	3272764	37%	3035599	93%	237165	7.2%	3589	233576

Table 2-1: Summary table of the data specifics for the each sample.

Chapter 3 :

A novel aminopepetidase acts in *Drosophila* testicular niche to regulate dedifferentiation of progenitor germ cells

Summary

Spermatogonia cells going through the proliferation stage have to decide to irrevocably commit to differentiate to eventually become sperm versus to dedifferentiate to replenish the stem cells lost during normal aging or tissue regeneration process. Despite the remarkable progress in understanding differentiation pathways in Drosophila stem cell systems, the molecular mechanism and factors involved in dedifferentiation process remains unclear. In our quest to discover new factors that regulate the niche functions, we have isolated a relatively homogenous niche sample containing hub and stem cells to compare with gonialblast cysts. Our analysis of these two samples has led to the identification of niche specific factors with potential critical functions in the niche.

One of the identified niche specific factors is an aminopeptidase encoded by the slamdance (sda) gene, whose function has never been shown previously in any stem cells system. Our characterization of this gene showed that loss of function of sda leads to dramatic abnormalities in the testis niche. Specifically, we observed niche architecture deterioration and stem cell loss from the niche. Our further studies showed that the GSCs loss in sda mutant is caused by both defects in cadherin-dependent maintenance of bona fide GSCs and dedifferentiation of spermatogonia cells to repopulate the niche during aging. Loss of function and gain of function assays also showed that Sda's function in the hub cells is both necessary and sufficient in promoting spermatogonia cells to undergo dedifferentiation. Using deletion and substitution of critical residue constructs, we are able to show that Sda activities in the niche are dependent on its catalytic domain.

Understanding how transit amplifying cells undergo dedifferentiation to revert back to plastic stem cells is critical in illuminating tissue homeostasis during aging and repair mechanism during tissue injury, especially in light of its potential application in regenerative medicine

therapies. My finding and characterization of this novel player in this dedifferentiation process offers an important advancement in understanding how a niche specific peptidase can influence differentiation versus dedifferentiation decision.

Introduction

Stem cells have the unique ability to self-renew and generate daughter cells which differentiate into distinct cell types (Knoblich, 2008; Morrison and Kimble, 2006). Adult stem cells are especially important to multicellular organisms since they are the only source to replenish their cells during turnover of regular cells due to aging or tissue injury. Adult stem cells normally reside in a specialized micro-environment called the niche that supplies the necessary factors needed to maintain stem cell identity and prevent precocious differentiation of these stem cells. (Losick et al., 2011; Morrison and Spradling, 2008). One very common way to ensure balance in the population of stem cells and its differentiating daughter cells is through physical attachment between stem cells and their niches. This way the niche can constantly supply the signaling molecules needed to maintain stem cells in a controlled short range manner. An exhaustive discussion of signaling molecules in Drosophila male stem cells niche is found in Chapter 1 of this thesis. Finally, this physical anchoring and selective supply of stem cells conducive microenvironment ensure that these stem cells will consistently undergo asymmetric cell divisions producing one stem cells and one differentiating cells at every division (Yamashita, 2010; Yamashita et al., 2010).

The *Drosophila* germline stem cell (GSC) niches have well-characterized physiological locations and cellular structures which allow precise identification of GSCs and differentiating cells using various markers and morphological criteria (Losick et al., 2011). In both female and male *Drosophila* GSC lineages, the daughter cells are displaced from the niche to undergo four round of mitosis followed by meiosis and terminal differentiation to eventually produce distinctive haploid gametes (Clarke and Fuller, 2006). Progenitor germ cells at the proliferative stage called spermatogonia have recently been shown to undergo dedifferentiation to reoccupy the niche and become GSC like cells.(Brawley and Matunis, 2004; Cheng et al., 2008; Kai and Spradling, 2004; Sheng et al., 2009; Sheng and Matunis, 2011). These GSC like cells have been

characterized to have molecular characteristic of GSCs such as physical attachment to the hub and the ability to undergo asymmetric division. This dedifferentiation has been shown to occur under physiological conditions such as aging(Cheng et al., 2008; Wong and Jones, 2012) and/or during recovery from genetically manipulated depletion of GSCs from the niche (Brawley and Matunis, 2004; Kai and Spradling, 2004; Sheng et al., 2009). Moreover, live imaging experiments have recently showed more clearly that these dedifferentiating spermatogonial cells move to the niche and make initial contacts with the niche(Sheng et al., 2009; Sheng and Matunis, 2011). Other studies have also shown factors in the differentiating spermatogonia that potentially affect its ability to dedifferentiate (Sheng et al., 2009; Yamashita, 2010). However, it is unclear whether and how niche promotes dedifferentiation (Sheng and Matunis, 2009). In this chapter, we showed that our unbiased analysis of transcripts enriched in the niche has identified an aminopeptidase that played a role in stem cells homeostasis.

Materials and Methods

Fly strains and husbandry

Fly stocks were raised using standard Bloomington medium at 25°C or 29°C as noted. The following fly stocks were used: *y*,*w* (as *wild-type* or WT), sda^{P} (Bloomington Stock Center BL-10344) as the enhancer trap line, $sda^{iso7.8}$ as a null allele(Zhang et al., 2002), *Df(3R)ED6235* that uncovers the *sda* gene region (BL-9878), *UAS-sda dsRNA* (Vienna Drosophila Research Center GD11680), *UAS-Dicer2* (Vienna Drosophila Research Center #V60008), *UAS-DE-Cadherin* (Sanson et al., 1996), *UAS-DN-Cadherin* (Iwai et al., 1997), $NCad^{M12}$ (BL-229) as an unprocessable DN-Cadherin mutation, dNc as an unprocessable DE-cadherin mutation (Oda and Tsukita, 1999), *UAS-upd* (Terry et al., 2006), *UAS-Socs36E* (from B. Callus, University of Western Australia, Perth, WA, Australia), 2×Stat reporter line (from E. Bach, New York

University School of Medicine, New York, NY, USA), *UAS-GFP.nls* (BL-4776), *UAS-Gal4* (BL 5939) and *hs-bam* (BL-24636). The Gal4 drivers are *upd-GAL4* (Leatherman and Dinardo, 2010), *c587-gal4* (from A. Spradling, Carnegie Institution Department of Embryology, Baltimore, MD, USA) and *nos-gal4* (Van Doren et al., 1998).

Temperature shift assay to knock down *sda* in adult flies

Flies with the *UAS-sda dsRNA*; *UAS-Dicer2* transgenes paired with different Gal4 drivers were raised using standard Bloomington medium at 18°C. For the negative control, flies were kept at 18°C throughout development until adulthood so that there is low or no knockdown of *sda*. For the positive control, flies were shifted to 29°C at L1 stage to maximize the knockdown effect. To specifically knock down *sda* in adult flies, newly eclosed males were shifted to 29 °C and kept at 29 °C for 12 days before analyzing the phenotypes.

Heat shock regime

Newly eclosed males with noted genotypes were collected and aged for 20 days with female flies in an 18°C incubator. Before heat shock, males were transferred to bottles that had been air dried for 24 hours. Bottles were submerged with all air area underneath water in a circulating 37°C water bath for 30 minutes at approximately 9 AM and 4 PM daily for 5 days, for a total of ten times of heat shock. Flies were placed in a 29°C incubator between heat shock treatments and returned to 18°C after the final heat shock. Flies were then allowed for a 7-day recovery at 18°C.

Generation of constructs and transgenic fly strains

Standard procedures were used for all subcloning procedures. Enzymes used for plasmid construction were obtained from New England Biolabs (Beverly, MA, USA), Promega Biotech (Madison, WI, USA), and Agilent (Santa Clara, CA, USA). The HA fragment was PCR amplified as a SpeI and BmtI flanked fragment from the *sa-3HA* plasmid with HA_F and HA_R primers (Chen et al., 2005).

1. *UAS-HA-sda*^{FL}: Using cDNA prepared from WT flies, the 5' *sda* cDNA P1 was PCR amplified as SpeI and NdeI flanked fragment with sda_F and sda_R2 primers, while the 3' *sda* cDNA P2 as NdeI and XbaI flanked fragment with sda_F2 and sda_R primers. The *sda* cDNA P1 and *sda* cDNA P2 were inserted using a three-way ligation into the pUASpB plasmid (a gift from Dr. Van Doren lab, Johns Hopkins University, Baltimore, MD, USA) cut with SpeI and XbaI. The HA fragment was subsequently inserted into *pUASpB-sda*^{FL} cut with SpeI and BmtI at the 5' end.

2. *UAS-HA-sda*$^{\Delta CAT}$: Using *sda*^{FL} as described above, the 5' *sda*$^{\Delta CAT}$ P1 was amplified as SpeI and NdeI flanked fragment with sda_F and sda_NCAT_R primers, while the 3' *sda*$^{\Delta CAT}$ P2 as NdeI and XbaI flanked fragment with sda_NCAT_F and sda_R primers. *sda*$^{\Delta CAT}$ P1 and P2 were inserted using a three-way ligation into the pUASpB plasmid cut with SpeI and XbaI. The HA fragment was subsequently inserted into *pUASpB-sda*$^{\Delta CAT}$ cut with SpeI and BmtI at the 5' end.

3. *UAS-HA-sda*$^{E\rightarrow A}$: To generate *sda*$^{E\rightarrow A}$, a fragment from *pUASpB-HA-sda*^{FL} was excised using BglII and XbaI and subcloned into pBluescript. Using QuikChange Site-Directed Mutagenesis Kit (Agilent Cat#200518), the E to A change was made with sda_EtoA_F and sda_EtoA_R primers according to manufacturer's protocols. The resulting fragment with E to A change was inserted back to *pUASpB-HA-sda*^{FL} cut with BglII and XbaI.

All plasmids (1 to 3) were introduced to attP40 flies using PhiC31 integrase-mediated germline transformation (Bestgene Inc., CA, USA).

Sequences of primers:

HA_F: 5'-AAAAACTAGTCAC ATT ATGGGCCGCATCTTT-3'

HA _R: 5'-AAAAGCTAGC CTGCTGCATACCGCTCTGAGC-3'

sda_F: 5'-AAAAACTAGTGCTAGCACCATGGAGGGCGGTTACGTCAACGAA-3'

sda_F2 : 5'-CCTCATCGGAACATATGCAGGTGGTGGCCG-3'

sda_R2: 5'-CGGCCACCACCTGCATATGTTCCGATGAGG-3'

sda_R: 5'-AAAATCTAGA CTCCCCCATTTCCAACTTCCATGTG-3'

sda_NCAT_F: 5'-AAAACATATGCCGGAGTTCCAGAGCATGGAC-3'

sda_NCAT_R: 5'-AAAACATATG AGTCCGGCACGGACACCAAATC-3'

sda_EtoA_F: 5'-GATTCGCTGCCATGGCCAACT GGGGACTC-3'

sda EtoA_R: 5'-GAGTCCCCAGTTGGCCATGGC AGCGAATC-3'

Immunostaining

Immunofluorescence staining was performed as described previously (Cheng et al., 2008). The primary antibodies used were as follows: mouse anti-ɣ-tubulin (1:100; Sigma T9026); mouse anti-Armadillo (1:100; DSHB N2 7A1 clone); mouse anti-α- Spectrin (1:20; DSHB 3A9 clone); EdU (Invitrogen C10350); Lysotracker (according to manufacturer recommendation, Invitrogen L7528); rabbit anti-Ser10-phosphorylated Histone H3 (1:200; Upstate 07-424); rat anti-Vasa (1:40; a gift from Dr. Allan Spradling, Carnegie Institution for Science); rabbit anti- Zfh1 (1:4000; a gift from Dr. Ruth Lehmann, Skirball Institute of Biomolecular Medicine); rabbit- anti

STAT [1:1000, a gift from Denise Montell (Silver et al., 2005)]; rabbit- anti active Caspase-3 [1:500, BD 559565], and mouse anti-β Galactosidase (1:5000; Promega z3781). Images were taken using a Zeiss Apotome microscope with a 63x oil immersion objective and processed using Adobe Photoshop software.

Quantifying *Drosophila* Lifetime

One hundred newly eclosed flies were collected (50 male and 50 female flies) and transferred to bottles that had been air dried for 24 hours. We added dry yeast to the bottles. Every 2 days, the flies were transferred into a new air-dried bottle with dry yeast. The dead flies at the bottom of the old bottle were counted. We repeated this process until all flies were dead.

Calculating mitotic index

We used EdU (Invitrogen C10350) incorporation assay and anti-phosphorylated histone 3 (H3S10P, Upstate 07-424) immunostaining to compute mitotic index as EdU- or PH3-positive GSCs/Total GSCs. Since GSC mitosis is sensitive to CO2 anesthetization, we dissected testes within 5 minutes of anesthetization followed by immediate fixation. The mitotic index results using both anti-PH3 immunostaining and EdU incorporation were based on two independent experiments, respectively.

Cell death assay for hub cells

The hub cells death were scored using Lysotracker (according to manufacturer recommendation, Invitrogen L7528) and Armadillo antibody (DSHB N2 7A1 clone). Armadillo and Lysotracker

double-positive cells were counted as dying hub cells. We also performed cell death assay using antibodies recognizing active Caspase-3 (1:500, BD 559565). Caspase-3 and Armadillo double-positive cells were counted as dying hub cells. We performed Lysotracker staining using third instar larvae (n=50) and pupae (n=50). We performed Caspase staining using third instar larvae (n=15) and pupae (n=10).

Lineage tracing experiment for hub cells

The *upd-Gal4*; *UAS-Gal4*; *UAS-GFP.nls* flies were used to permanently label hub cells with nuclear GFP (Le Bras and Van Doren, 2006). We then performed immunostaining experiments using anti-Arm and anti-Vasa. Specifically, GFP-positive cells outside the Arm-positive hub area were considered as cells leaving the hub. In total, we examined 50 testes from D1-3 *sda/Df* adult flies (n=50).

Calculation of 95% confidence interval (CI)

95% CI = Mean ± (1.96 x SE); SE: standard error.

P-value calculation and explanation

We used student *t*-test and Fisher exact test for P-value calculation. Student *t*-test was used when the experimental data is on continuous scale. Fisher exact test was used when a null hypothesis was tested and the experimental data had only two possibilities (e.g. equal *vs.* unequal).

Results

Characterization of Sda mutant testis

In *Drosophila* testis, GSCs associate with two types of somatic cells: hub cells and cyst stem cells (CySCs) (Fig. 3-1A). In the niche of *slamdance* (*sda*) mutant testes, substantial changes were detected for all three cell types (Fig. 3-1, Fig. 3-2). Compared to 30-day old (D30) *wild-type* (WT) testes (Fig. 3-1B-B'), hub cells in the *sda* mutant were reduced (Fig. 3-1C-C') even though no hub cell was found to undergo cell death or trans-differentiation to other cell types (Materials and Methods). We also found that the numbers of both GSCs (Fig. 3-1B-C) and Zfh-1-positive early cyst cells including CySCs (Issigonis et al., 2009; Leatherman and Dinardo, 2008) (Fig. 3-1D-E) significantly decreased. Similar results were obtained in *sda* maternal and zygotic mutants, suggesting that the defects at later developmental stages did not result from maternal contribution. To determine whether loss of both hub cells and stem cells in *sda* mutant occurred at the time of niche establishment in embryogenesis (Le Bras and Van Doren, 2006) or later during niche maintenance, a time-course experiment was performed using *sda* mutant and WT testes across developmental stages, from the first instar larvae (L1) to D30 adult males. For all three cell types, no substantial difference was observed at the early developmental stages (e.g., L1 for hub cells and GSCs, L3 for Zfh-1-positive cells, Fig.3-1F-H), suggesting maintenance rather than establishment defects. Notably, Sda promotes GSC expansion from L2 to L3 stage because GSCs increase 1.7-fold in WT testes and 1.3-fold in *sda* mutant testes, respectively (Fig.3-1G).

Characterization of Sda expression pattern and functionality

We next asked in which cell type Sda is functional to maintain the normal testicular niche. Using RNA-seq analysis we found high level of *sda* transcript in the niche sample

including hub cells, GSCs and CySCs (Fig. 3-3A). We further showed that expression of *sda* is mainly detectable in hub cells using an enhancer trap line (Zhang et al., 2002) (Fig. 3-4 A-A'). Consistent with its expression pattern, the endogenous function of Sda was required exclusively in hub cells based on two complementary assays. First, the *sda* RNAi transgene (*UAS-ds sda*) driven by a hub-specific *upd-Gal4* driver (Leatherman and Dinardo, 2010) resulted in phenotypes resembling the *sda* mutant (Fig. 3-4B-D, Fig. 3-5A). Consistent with the maintenance function of Sda shown previously (Fig. 3-1), a temperature shift assay (Eliazer et al., 2011) that specifically knocked down *sda* in adult flies was sufficient to recapitulate all *sda* mutant phenotypes in the testicular niche (Fig. 3-6A). Second, the *sda* mutant phenotypes were fully rescued by driving an HA-tagged full-length *sda* cDNA (*HA-sda*^{FL}) using the same *upd-Gal4* driver (Fig. 3-4 E-F, I). By contrast, the use of a germline-specific *nanos-Gal4* driver (Van Doren et al., 1998) or a cyst cell-specific *c587-Gal4* driver (Manseau et al., 1997) showed much less effect in either the knockdown (Fig. 3-4D) or the rescue (Fig. 3-6 B) experiment. In summary, these results demonstrate that Sda acts specifically in hub cells to maintain a normal testis niche.

Sda has been predicted to be a single transmembrane protein that is homologous to the mammalian zinc-dependent aminopeptidase N (APN) enzyme (Zhang et al., 2002). Consistently, we found that *HA-sda*^{FL} is localized to the plasma membrane when expressed in hub cells (Fig. 3-6 C). The identity and similarity between Sda and human APN is 33% and 51%, respectively (31% identity and 50% similarity between Sda and mouse APN), while their catalytic domains are as high as 84% identical (Zhang et al., 2002). Although *sda* mutant was shown to have neuronal defects (Zhang et al., 2002), we did not find significant shorter lifespan of *sda* males compared to control males (Fig. 3-6 F). To study whether Sda acts as an APN in the testicular niche, two *sda* cDNAs encoding putative enzymatically inactive forms were generated: an *HA-sda*$^{\Delta CAT}$ with a truncation at the predicted zinc-binding and APN domain; and an *HA-sda*$^{E\rightarrow A}$ with a point mutation at the Glu (E) residue in the AAMEN domain, which is a critical site for

substrate recognition (Luciani et al., 1998; Zhang et al., 2002) (Fig. 3-4 E). Neither transgene rescued the *sda* mutant phenotype when driven by *upd-Gal4* (Fig. 3-4 G-I, Fig. 3-6 D-E, G, Fig. 3-5A), suggesting that the catalytic domain is required for normal function of Sda in the testis niche.

Sda is necessary and sufficient to promote spermatogonia dedifferentiation

We next investigated mechanisms underlying the substantial GSC loss in *sda* mutant (Fig. 3-1, Fig. 3-2). During aging hub cells decrease (Wallenfang et al., 2006) but GSCs are maintained (Cheng et al., 2008) or decreased only slightly (Boyle et al., 2007). Maintenance of GSC number during homeostasis is contributed to both self-renewal (Sheng and Matunis, 2011; Yamashita et al., 2003) of *bona fide* GSCs established during embryogenesis (Le Bras and Van Doren, 2006) and spermatogonial dedifferentiation (Cheng et al., 2008; Sheng and Matunis, 2011). We first checked mitotic activity of GSCs using both an anti-PH3 (Lim and Fuller, 2012) immunostaining and a pulse EdU incorporation assay (Insco et al., 2009). The percentage of PH3-positive GSCs for *sda* mutant (2.97%, n=101) was higher than that of WT testes (1.50%, n=334), so was the percentage of EdU-positive GSCs [*sda* mutant: 21% (n=125); WT: 17% (n=526)]. Therefore loss of GSCs in *sda* mutant cannot be attributed to decreased GSC. We next investigated whether spermatogonial cells fail to undergo dedifferentiation in *sda* mutant testes. It has been shown that dedifferentiated spermatogonial cells tended to have misoriented centrosomes (Cheng et al., 2008). Indeed, in *sda* mutant testes we found a significant deprivation of GSCs with misoriented centrosomes (Fig.3-7 A-C), suggesting a potential dedifferentiation defect in *sda* mutant testes. In addition to misoriented centrosomes, it has been reported that transient disintegrating fusome remnants are detectable in dedifferentiated spermatogonial cells next to hub cells (Brawley and Matunis, 2004; Cheng et al., 2008; Inaba et al., 2010; Sheng et al., 2009). Consistent with dedifferentiation defects, GSCs with disintegrating fusomes were

significantly reduced in *sda* mutant testes (Fig. 3-7 D-E, 3-7 G). Taken together, these data suggested that Sda is necessary for dedifferentiation of spermatogonial cells.

Next, we investigated whether overexpression of Sda is sufficient to promote spermatogonial dedifferentiation during homeostasis. We found that overexpression of full-length *sda* in hub cells led to ~4-fold more GSCs with disintegrating fusome remnants compared to that in the control testes (either *upd-Gal4* or *UAS-sda*^{FL} only, or *upd> sda*^{FL}; *sda*) under the same condition, suggesting that Sda is sufficient to promote spermatogonial dedifferentiation (Fig.3-7 F-G). In addition, such an effect of Sda overexpression in promoting spermatogonial dedifferentiation could be detected as early as D1 and lasted throughout adulthood (Fig. 3-5 B). However, despite increased spermatogonial dedifferentiation, overexpression of *sda* did not lead to an increase of overall GSC number [10.8±0.9s.d. GSCs in *upd>sda*^{FL} testes (n=65), versus 10.5±1.7s.d. GSCs in WT testes (n=33), P>0.1], suggesting that dedifferentiated spermatogonial cells may outcompete *bona fide* GSCs at the niche as shown previously in the *Drosophila* ovary (Jin et al., 2008).

Sda is required for testis niche to home dediferentiation during tissue regeneration

To further investigate whether Sda is required for dedifferentiation when GSCs are depleted by genetic manipulations, a differentiation factor encoded by *bag-of-marbles* (*bam*) was ectopically expressed in all cells including GSCs (Sheng et al., 2009) using a rigorous heat shock regime (Fig. 3-8A). As a result, GSCs differentiated and left the niche in the control and *sda* mutant testes (Fig. 3-8B, 3-8D), leaving ~0.15 GSC per testis (Fig. 3-8F) and ~85.3% of testes with zero GSC (Fig. 3-8G). Upon subsequent recovery, 58.7% of control testes regained GSCs (Fig. 3-8C, 3-8G) through dedifferentiation of spermatogonial cells, leading to an average of 7.2 GSCs per testis (Fig. 3-8F). By contrast, 95.7% *sda* mutant testes had zero GSC after the same recovery time, suggesting that spermatogonial cells failed to dedifferentiate (Fig. 3-8E-G).

Furthermore, we found that only full-length sda^{FL} but not the $sda^{\square CAT}$ or the $sda^{E\rightarrow A}$ could rescue the dedifferentiation defect in *hs-bam*; *sda* mutant males (Fig. 3-8F-G), suggesting that the catalytic activity of Sda is required for spermatogonial dedifferentiation during tissue regeneration.

Cadherins molecules are potential Sda substrates

To address how Sda promotes spermatogonial dedifferentiation, we took a candidate gene approach to investigate cell-cell adhesion molecules (Inaba et al., 2010) and signaling pathway components (Sheng et al., 2009) for their interactions with Sda. It has been shown that mis-expression of a dominant negative form of *Drosophila* E-cadherin homolog (DE-cadherin, E-cad) in germ cells reduces dedifferentiation (Inaba et al., 2010). In addition, ectopic expression of *suppressor of cytokine signaling 36E* (*Socs36E*), an inhibitor of the JAK-STAT (Janus kinase and signal transducer and activator of transcription) signaling pathway, was reported to reduce spermatogonial cell dedifferentiation (Sheng et al., 2009). However, in both examples gene expression is manipulated in germ cells but not in niche cells.

We found that removal one copy of Cadherins made the *sda* loss-of-function phenotype more severe (Fig. 3-9A). Furthermore, expression of either *E-Cad* (Fig. 3-9 B-C) or *Drosophila* N-cadherin homolog (DN-cadherin, N-cad) in hub cells (Fig. 3-9 D), but not in germ cells (Fig. 3-9 E), rescued most of the *sda* mutant phenotypes. Because Sda is predicted to be a peptidase and maturation of both E-Cad and N-Cad requires protein cleavage (Iwai et al., 1997; Oda and Tsukita, 1999), we examined whether Sda processes Cadherins post-transcriptionally. Indeed, although the overall transcript level of *E-Cad* was comparable and even slightly higher in *sda* mutant testes than that in WT control (Fig. 3-9F), the mature E-Cad protein level decreased (Fig. 3-9G). Finally, using either an $NCad^{M12}$ allele that cannot be processed by peptidase (Iwai et al.,

1997) or driving an *E-Cad* with the cleavage site mutated [i.e. unprocessable E-Cad, (Oda and Tsukita, 1999)] resulted in phenotypes similar to the *sda* mutant testes (Fig. 3-9H). Together, these data suggested that Cadherins are important downstream targets of Sda, although they may or may not be the direct substrates of Sda. On the other hand, Sda may have multiple substrates *in vivo* for different functions. Indeed, although both hub cells and Zfh-1-positive cells were fully rescued by overexpression of Cadherins (Fig. 3-9D), the GSC number was only recovered to ~84% of that in WT control. Further analysis indicated that the dedifferentiation defects in *sda* mutant testes were not rescued (Fig. 3-10), suggesting that dedifferentiation defects of spermatogonial cells can be separated from loss of hub cells and CySCs.

Sda promotes dedifferentiation independent of JAK-STAT signaling

To explore whether spermatogonial dedifferentiation defects in *sda* mutant testes result from compromised JAK-STAT signaling in germ cells, we examined the expression pattern of Stat92E in *sda* mutant testes. The GSCs in *sda* mutant testes were enriched with Stat92E (Fig. 3-11A-C'), suggesting that JAK-STAT signaling was properly received by GSCs in *sda* mutant testes. Furthermore, overexpression of the JAK-STAT ligand Upd in hub cells (Boyle et al., 2007; Toledano et al., 2012) was not sufficient to rescue GSC loss in the *sda* mutant (Fig. 3-11D), suggesting that Sda promotes dedifferentiation independent of JAK-STAT signaling.

Discussion

Despite remarkable progress made in understanding stem cell differentiation pathway, the molecular mechanisms governing the dedifferentiation process remain unclear. Here our data demonstrate that a novel aminopeptidase acts in the *Drosophila* testis niche, whose function is both necessary and sufficient to promote spermatogonial dedifferentiation for GSC homeostasis and regeneration (Fig. 3-12). Interestingly, Sda also showed specific expression in female GSC niche (Fig. 3-3B-B'), suggesting that it could act as a niche-specific factor in multiple stem cell systems. In mouse testes, differentiating spermatogonial cells can also undergo dedifferentiation to become germinal stem cells in a process that shares many cellular commonalities with *Drosophila* spermatogonial dedifferentiation (Barroca et al., 2009; Nakagawa et al., 2010). It is very likely that the molecular mechanisms we learn from studying *Drosophila* male GSC lineage will be useful in understanding parallel pathways in mammals (Chen, 2008). Understanding molecular mechanisms underlying progenitor cell dedifferentiation will greatly assist in the application of dedifferentiated cells from the same lineage to repopulate the endogenous niche and function like *bona fide* stem cells for tissue regeneration..

With this finding, there are still many future experiments to be done such as finding additional Sda substrate through: 1. Candidate approach such as genetic screen on factors that have been shown to be important in niche's function or 2.Unbiased approach such as the TAIL (Terminal Amine Isotopic Labeling) (Kleifeld et al., 2010) or yeast 2 hybrid method (Murali et al., 2011). Another interesting question would be to discover mouse or human equivalence of Sda which we can potentially test by trying to rescue mutant testis with mammalian Sda expression.

Finally, we conclude that understanding molecular mechanisms underlying progenitor cell dedifferentiation will greatly assist in applying dedifferentiated cells from the same lineage to repopulate the endogenous niche and function like *bona fide* stem cells for tissue regeneration.

Our findings provide the first initial steps into understanding this exciting field of regenerative medicine.

Figures

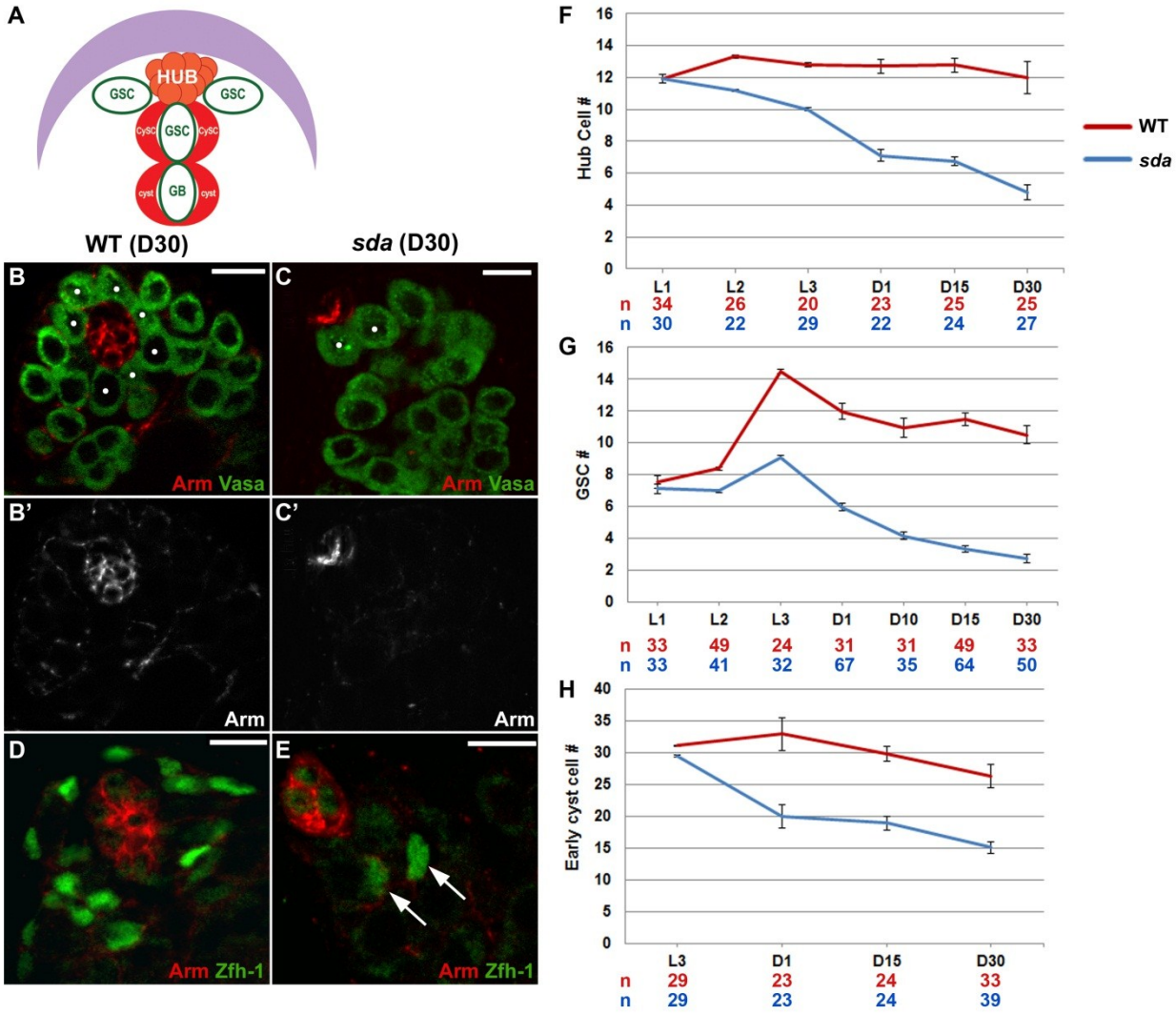


Figure 3-1: Sda is required for maintaining stem cells and hub cells in the testicular niche. (A) A schematic diagram outlines the *Drosophila* testicular niche (the purple crescent outlines the testis apical tip). **(B-E)** Immunostaining of testes from D30 WT (**B**, **B'**, **D**) and *sda/Df* (*sda*) mutant (**C**, **C'**, **E**) males using anti-Vasa (germ cells), anti-Armadillo (Arm) (hub cells) and anti-Zfh-1 (CySCs and early stage cyst cells); dots in (**B-C**) label GSCs, arrows in (**E**) point to the two Zfh-1-positive cells. Scale bar: 10μm.

(**F-H**) Quantification of hub cells (**F**), GSCs (**G**), and Zfh-1-positive cells (**H**) in testes from WT and *sda* mutant males at different developmental stages (L1, L2 and L3: first, second and third instar larvae; D1, D15 and D30: 1-day-, 15-day- and 30-day-old adult males.). Error bar: 95% Confidence Interval (CI). P value calculated by one-tailed *t*-test: no significant (n.s.) for L1 and <0.0001 for L2-D30 in (**F**); n.s. for L1, <0.01 for L2, <0.0001 for L3-D30 in (**G**); =0.01 for L3, <0.0001 for D1-D30 in (**H**).

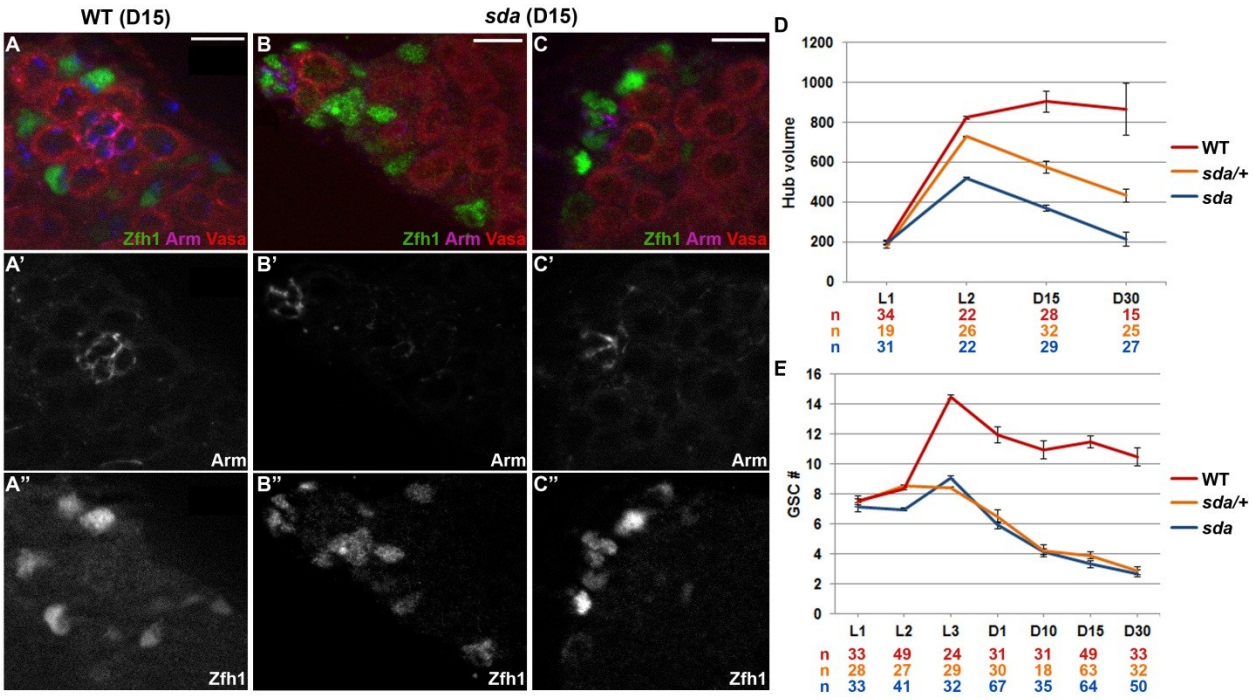


Figure 3-2: Mutant phenotype in testes from D15 *sda* mutant males (A-C") and *sda/+* males at different developmental stages (D-E). (**A-C"**) Immunostaining using anti-Vasa, anti-Arm, and anti-Zfh-1, in testes from D15 WT (**A**, **A'**, **A"**) or *sda/Df* (**B**, **B'**, **B"**; **C**, **C'**, **C"**) mutant males. Scale bar: 10μm. (**D-E**) Quantification for hub volume across developmental stage L1 to D30 for WT, *sda/+*, and *sda* (**D**). Quantification for GSC number across developmental stage L1 to D30 for WT, *sda/+*, and *sda* (**E**).

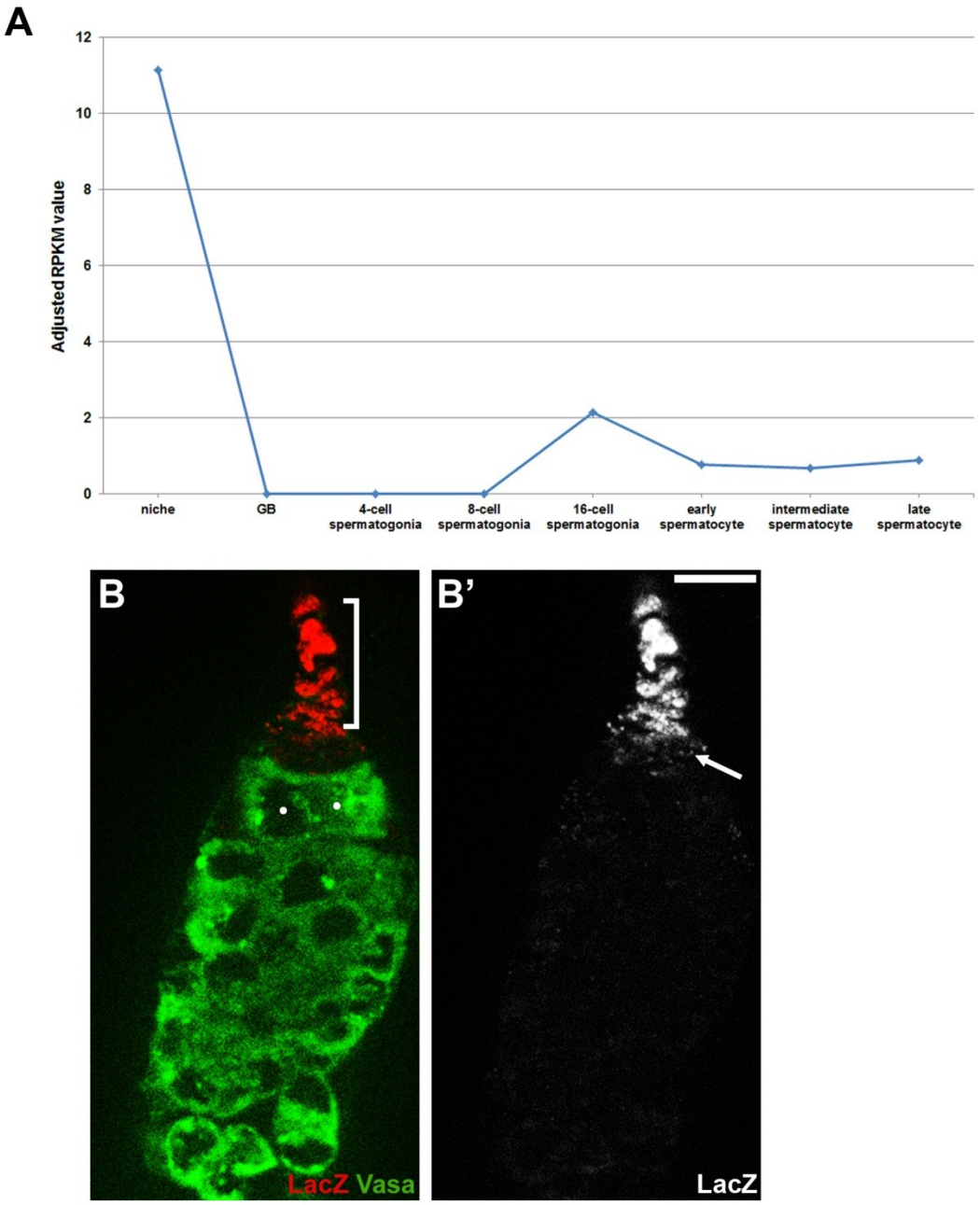


Figure 3-3: Expression pattern of *sda* in both germline cysts isolated from WT testes and in female ovary using an enhancer trap line. (A) Adjusted RPKM (reads per kilobase transcript per million of mapped reads) of *sda* transcript level in isolated niche sample (including hub cells, GSCs and CySCs), in comparison with other stage germ cell cysts, such as gonialblast, 4-cell spermatogonia, 8-cell spermatogonia, 16-cell

spermatogonia, as well as early, intermediate and late spermatocyte samples using RNA-seq analysis. The *sda* is highly transcribed in the niche sample compared to other stage germ cell cysts. (**B-B'**) Immunostaining using anti-Vasa and anti-LacZ in an ovariole from an *sda* enhancer trap line in which a *lacZ* reporter is inserted near the endogenous *sda* gene (Zhang et al., 2002). The LacZ staining is shown separately in (**B'**). Bracket indicates terminal filament with high LacZ and the arrow points to cap cells with low LacZ, both cell types contribute to female GSC niche. Scale bar: 10μm.

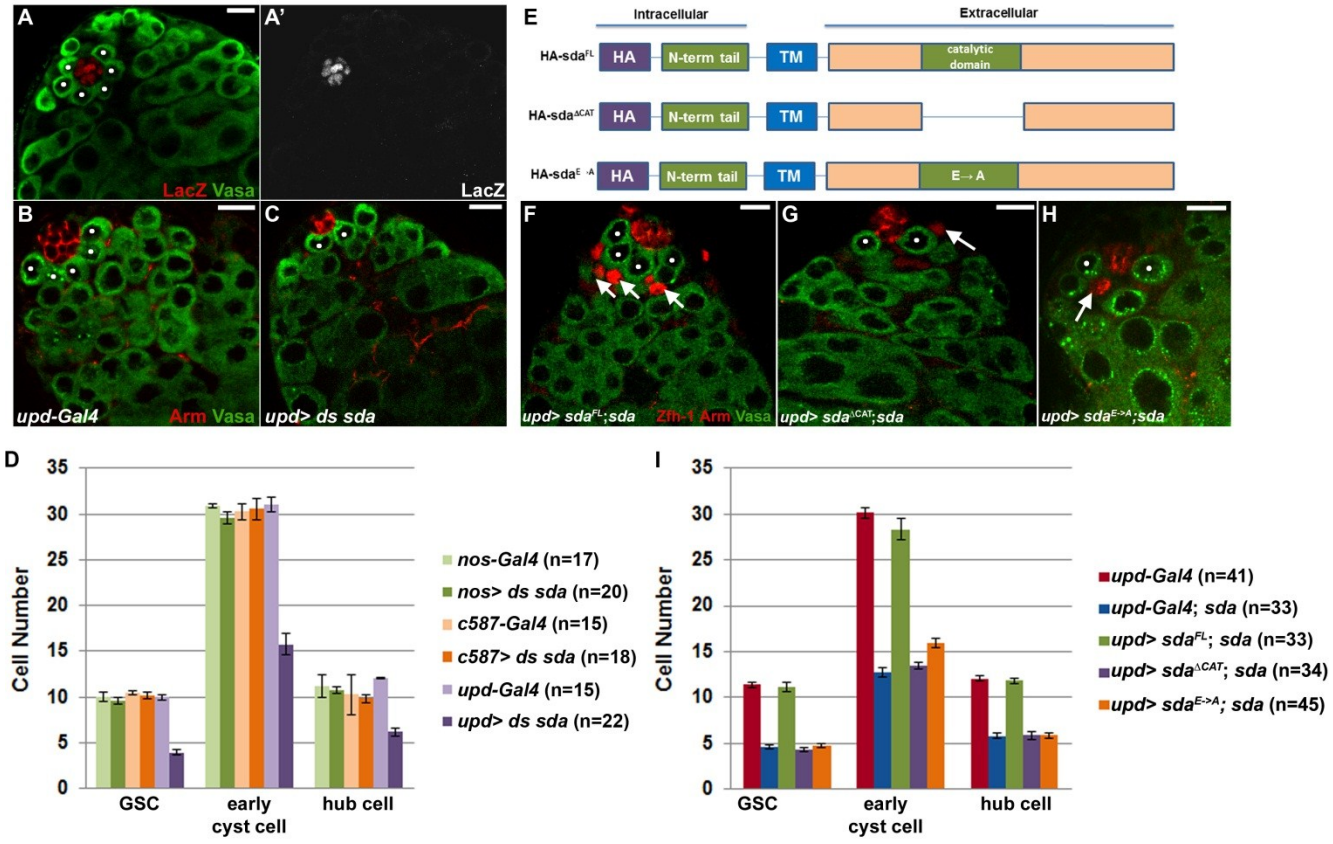


Figure 3-4: Sda acts specifically in hub cells to maintain niche structure and stem cells. (A) Immunostaining using anti-Vasa and anti-LacZ in testes from an *sda* enhancer trap line in which a *lacZ* reporter is inserted near the endogenous *sda* gene (Zhang et al., 2002). The LacZ staining is shown separately in (**A'**). (**B-C**) Testes from *upd-Gal4* (**B**)

and *upd-Gal4*; *UAS-ds sda*; *UAS-dcr2* (**C**) males, stained with anti-Vasa and anti-Arm, dots label GSCs. (**D**) Quantification of GSCs, Zfh-1-positive cells and hub cells in testes from *nos-Gal4*, *c587-Gal4* and *upd-Gal4* by themselves as controls and crossed to the *UAS-ds sda*; *UAS-dcr2* background. (**E**) Structure of *HA-sda*^{FL}, *HA-sda*$^{\Delta CAT}$ and *HA-sda*$^{E\rightarrow A}$, TM: transmembrane domain. **(F-H)** Testes from *upd-Gal4*; *UAS-HA-sda*^{FL}; *sda* (**F**), *upd-Gal4*; *UAS-HA-sda*$^{\Delta CAT}$; *sda* (**G**), and *upd-Gal4*; *UAS-HA-sda*$^{E\rightarrow A}$; *sda* (**H**) males stained with anti-Vasa, anti-Zfh-1 and anti-Arm, arrows point to Zfh-1-positive cells. Scale bar: 10μm. (**I**) Quantification of GSCs, Zfh-1-positive cells and hub cells. All quantification data were obtained using D15 males. Error bar: 95% CI.

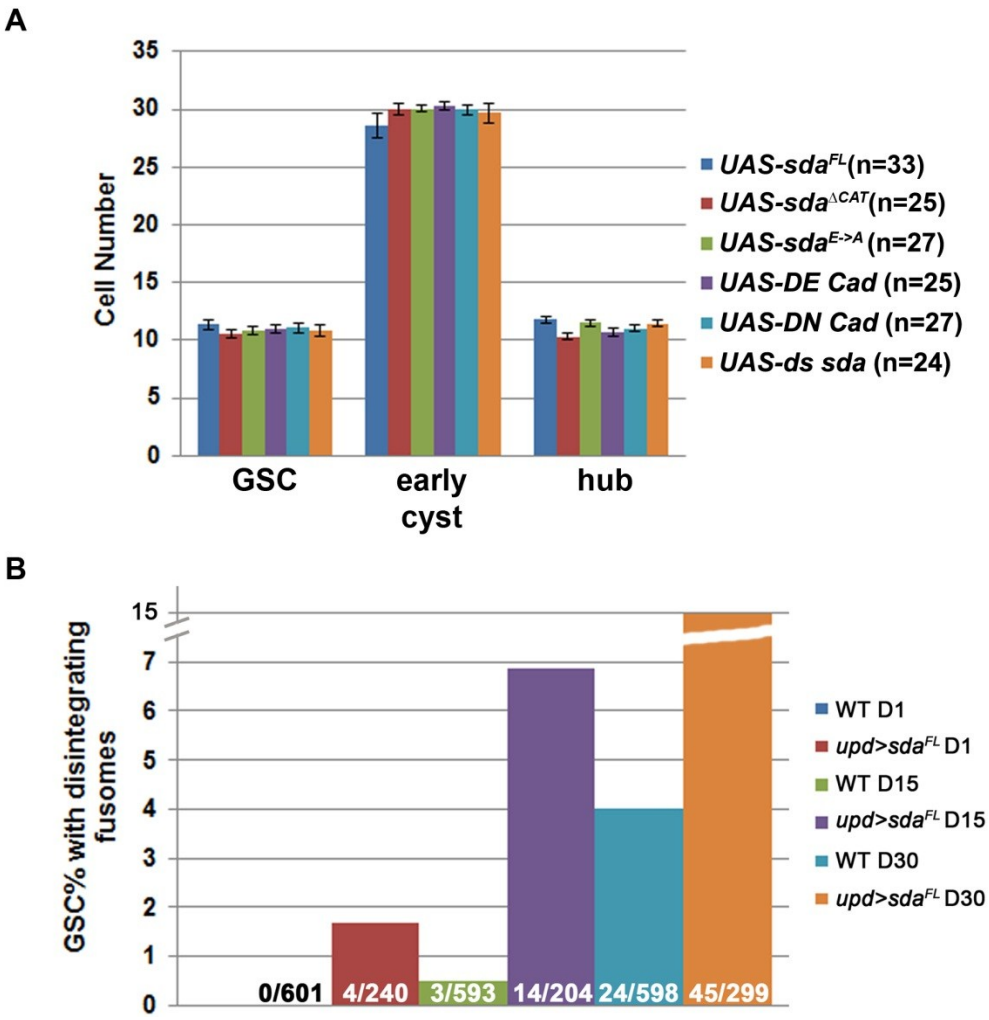


Figure 3-5: Quantification of different cell types in all *UAS* controls and overexpression of *sda* promotes spermatogonial dedifferentiation throughout adulthood. (A) Quantification of GSC, Zfh-1-positive cell and hub cell number in testes from all *UAS* controls which were outcrossed and aged at D15. (**B**) Overexpression of Sda using *upd-Gal4*; *UAS-HA-sda*^{FL} led to increased cysts with degenerating fusomes near the hub, in testes from D1, D15 and D30 adult males.

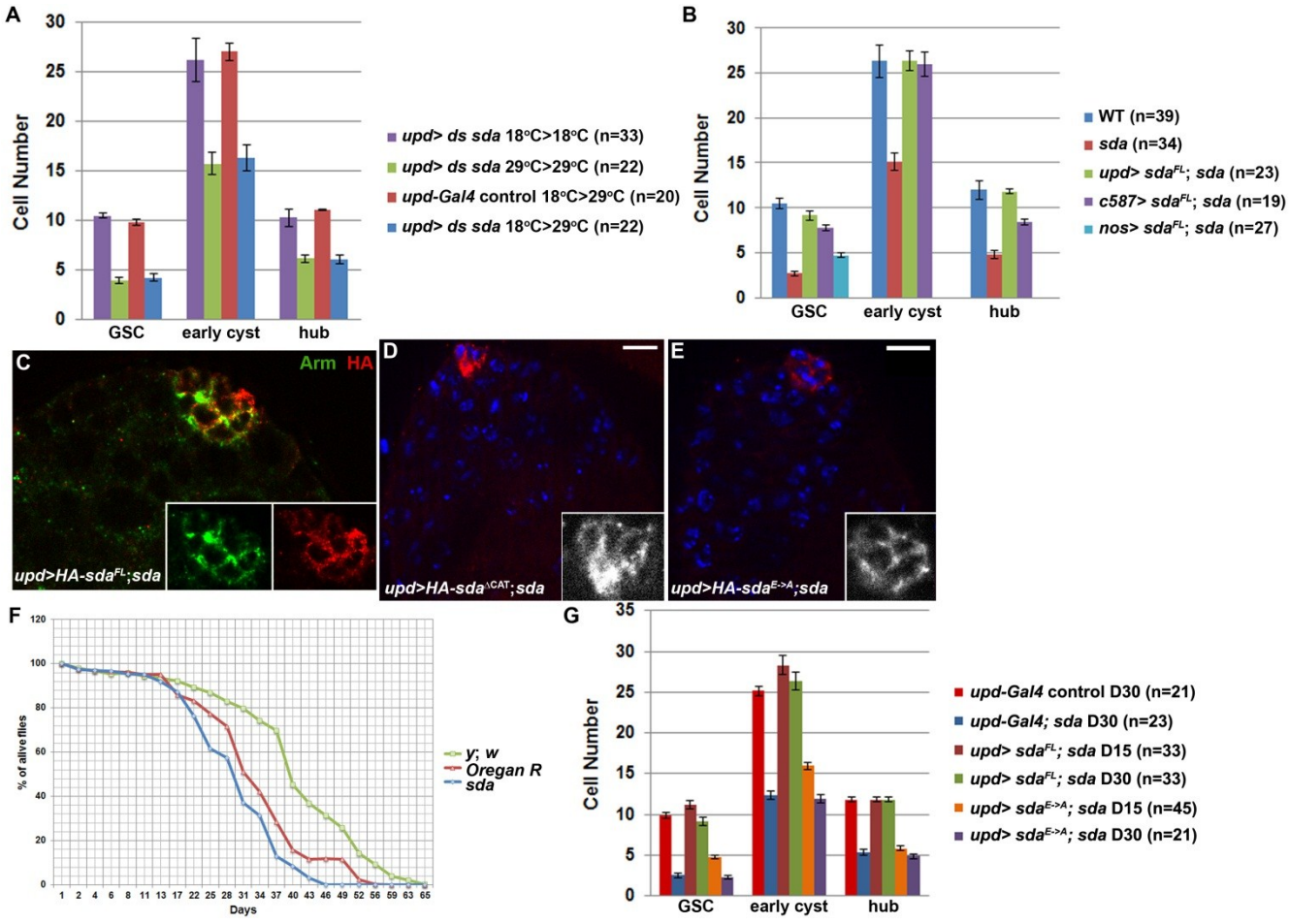


Figure 3-6: Cell-type-specific rescue using full-length *sda* cDNA and subcellular localization of different forms of Sda protein. (A) Quantification of GSC, Zfh-1-

positive cell and hub cell number in testes from *upd-Gal4* control and *upd-Gal4*; *UAS-ds sda*; *UAS-dcr2* raised under the designated temperature. **(B)** Quantification of GSC, Zfh-1-positive cell and hub cell number in testes from WT, or *sda*, or *upd-Gal4*; *UAS-HA-sda^{FL}*; *sda*, or *c587-Gal4*; *UAS-HA-sda^{FL}*; *sda*, or *nos-Gal4*; *UAS-HA-sda^{FL}*; *sda* males. (**C-E**) Immunostaining using anti-HA (red) and anti-Arm (green) in testes from *upd-Gal4*; *UAS-HA-sda^{FL}*; *sda* (**C**), *upd-Gal4*; *UAS-HA-sda$^{\Delta CAT}$*; *sda* (**D**), and *upd-Gal4*; *UAS-HA-sda$^{E\rightarrow A}$*; *sda* (**E**) males, blue: DAPI, inset: enlarged anti-HA staining in hub cells. (**F**) Longevity assay for two WT strains: *y*; *w* and *Oregan R*, and *sda/Df* line, based on five independent experiments. The lifespan difference between *sda/Df* males and WT males is smaller than the variation between the two WT strains. (**G**) Quantification of GSC, Zfh-1-positive cell and hub cell number in testes from *upd-Gal4* control, or *upd-Gal4; sda/Df*, or *upd-Gal4*; *UAS-HA-sda^{FL}*; *sda/Df*, or *upd-Gal4*; *UAS-HA-sda$^{E\rightarrow A}$*; *sda/Df* males at designated ages.

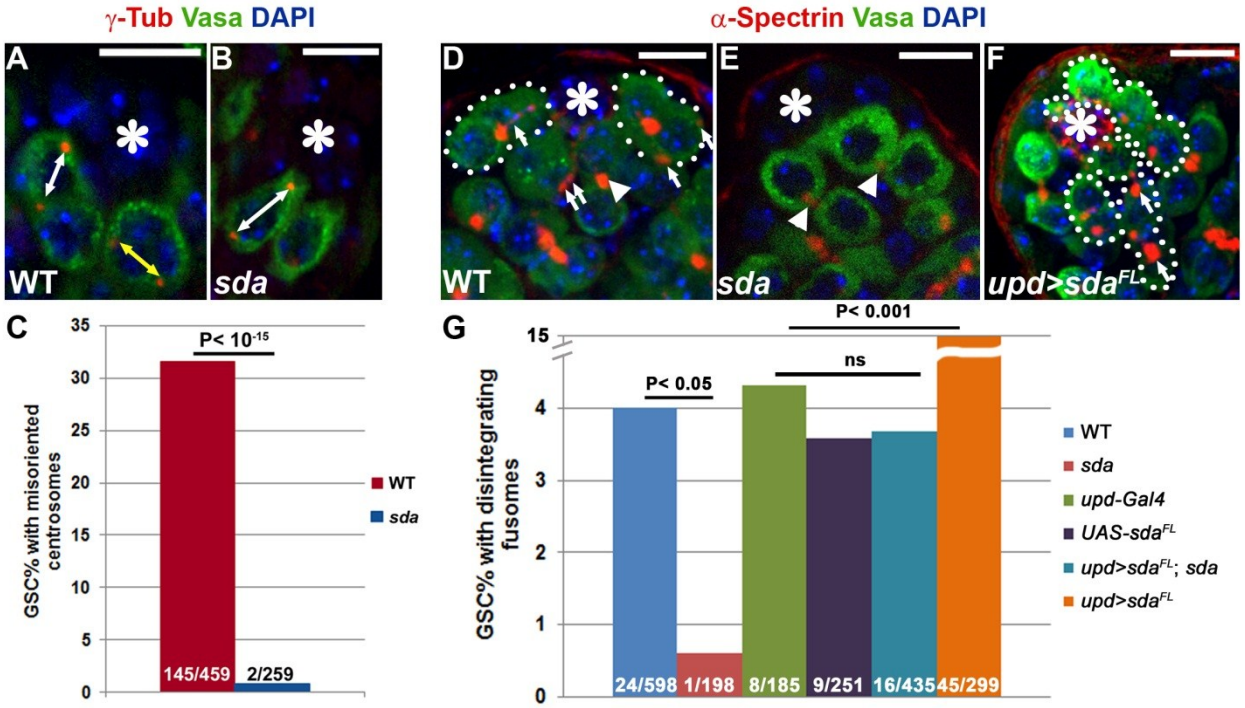


Figure 3-7: Sda is both necessary and sufficient to promote spermatogonial dedifferentiation during homeostasis. (**A-B**) Centrosomes (anti-γ-Tubulin) are misoriented in D30 WT (**A**, yellow double-arrowed line), but always show proper orientation in *sda* mutant testes (**B**, white double-arrowed line). (**C**) Percentage of GSCs with misoriented centrosomes in D30 WT and *sda* mutant testes. (**D**) Germ cell cysts with multiple spectrosomes (arrows, anti-α-Spectrin), or cysts with degenerating fusomes (circled by dotted lines) in D30 WT testes. (**E**) Only GSC-GB (gonialblast: daughter cell of GSC which will undergo proliferation and differentiation) pairs connected by a single spectrosome were detected in *sda* testes (arrowheads, also in **D**). (**F**) Overexpression of *sda* using *upd-Gal4*; *UAS-HA-sda*^{FL} led to increased cysts with degenerating fusomes (circled by dotted lines) next to the hub. Asterisks: hub. (**G**) Quantification of percentage of GSC with disintegrating fusomes; P-value calculated using Fisher's Exact Test. Scale bar: 10μm.

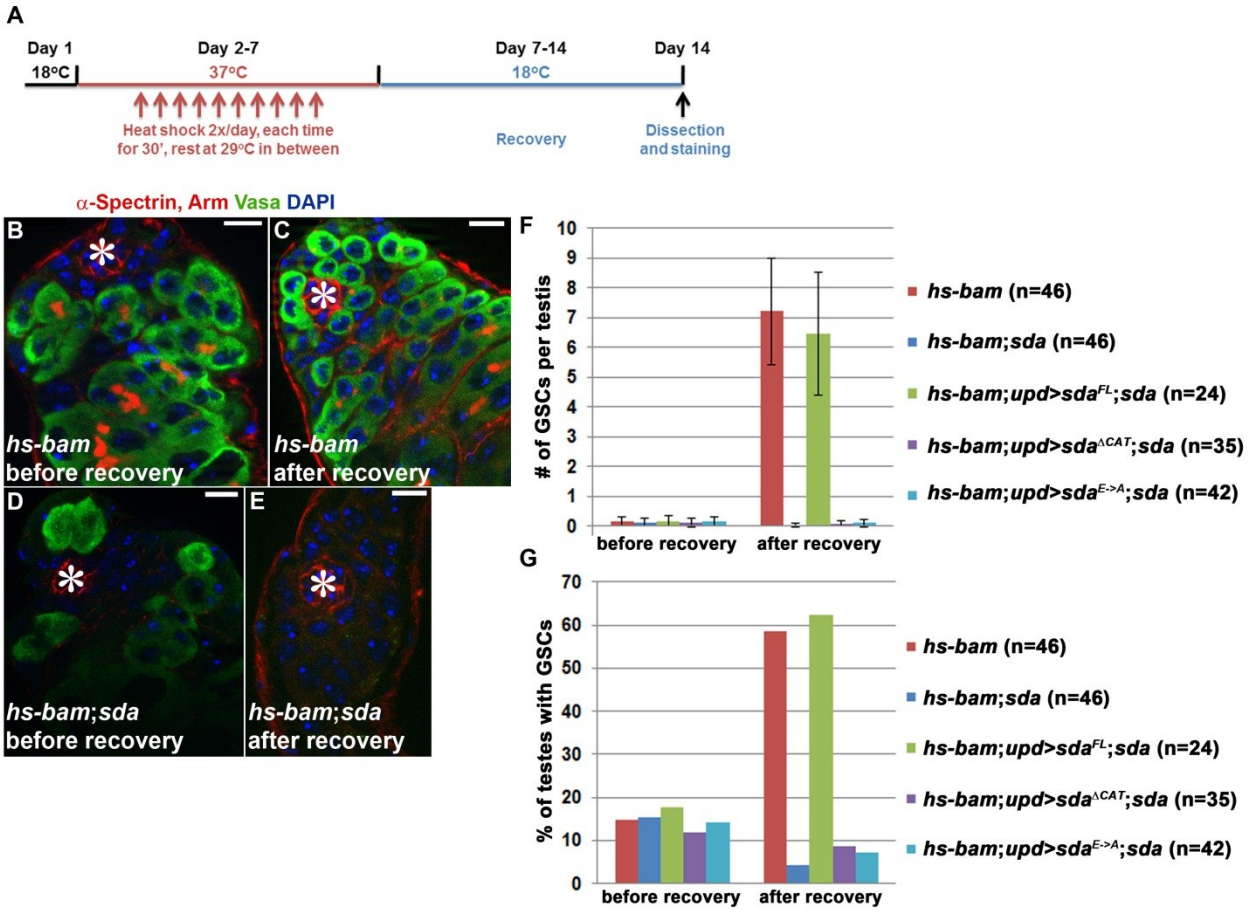


Figure 3-8: Sda is required for the testicular niche to home dedifferentiated spermatogonial cells during tissue regeneration. (A) Heat shock regime [modified from (Sheng et al., 2009)]. (**B-C**) Testes from *hs-bam* control males after heat shock treatment before recovery (**B**) and after recovery (**C**). (**D-E**) Testes from *hs-bam*; *sda* males after heat shock treatment before recovery (**D**) and after recovery (**E**). (**F-G**) Quantification of recovery efficiency, presented as the average number of GSCs (**F**) and the percentage of testes containing at least one GSC (**G**) in males with the corresponding genotype.

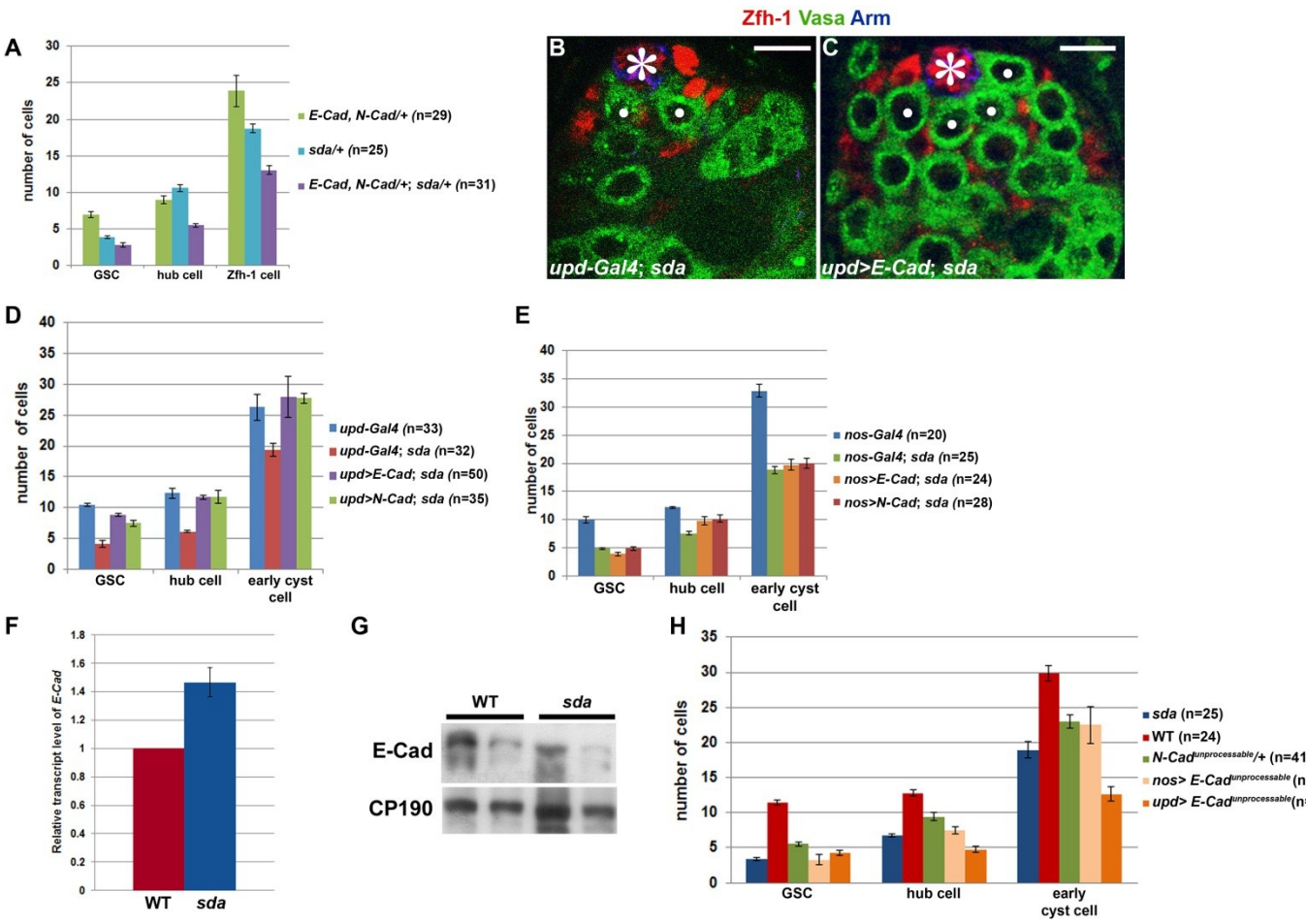


Figure 3-9: Sda processes Cadherin molecules post-transcriptionally to maintain the testicular niche. (A) Quantification of GSC, Zfh-1-positive cell and hub cell number in testes from *E-Cad*, *N-Cad/+*, or *sda/+*, or *E-Cad*, *N-Cad/+*; *sda/+* males. (**B-C**) Overexpression of E-Cad in hub cells using *upd-Gal4*; *UAS-E-Cad*; *sda* (**C**) rescued most *sda* mutant (**B**) phenotype in testicular niche. Scale bar: 10μm. (**D**) Quantification of all three cell types in the testicular niche from *upd-Gal4*, or *upd-Gal4*; *sda*, or *upd-Gal4*; *UAS-E-Cad*; *sda*, or *upd-Gal4*; *UAS-N-Cad*; *sda* D15 males. (**E**) Quantification of all three cell types in the testicular niche from *nos-Gal4* or *nos-Gal4*; *sda/Df* or *nos-Gal4*; *UAS-E-Cad*; *sda/Df* or *nos-Gal4*; *UAS-N-Cad*; *sda/Df* D15 males. Error bar: 95% CI. (**F**) Quantitative RT-PCR to measure *E-Cad* transcript levels in WT and *sda* mutant testes

based on three independent experiments. (**G**) Immunoblot to measure mature E-Cad protein levels in WT and *sda* mutant testes, CP190 is used as a loading control. **(H)** Quantification of GSC, Zfh-1-positive cell and hub cell number in testes from *sda/Df*, or WT, or $CadN^{M12}$ ($N\text{-}Cad^{unprocessable}$/+), or *nos-Gal4*; *dNc* ($nos>E\text{-}Cad^{unprocessable}$), or *upd-Gal4*; *dNc* ($upd>E\text{-}Cad^{unprocessable}$) males.

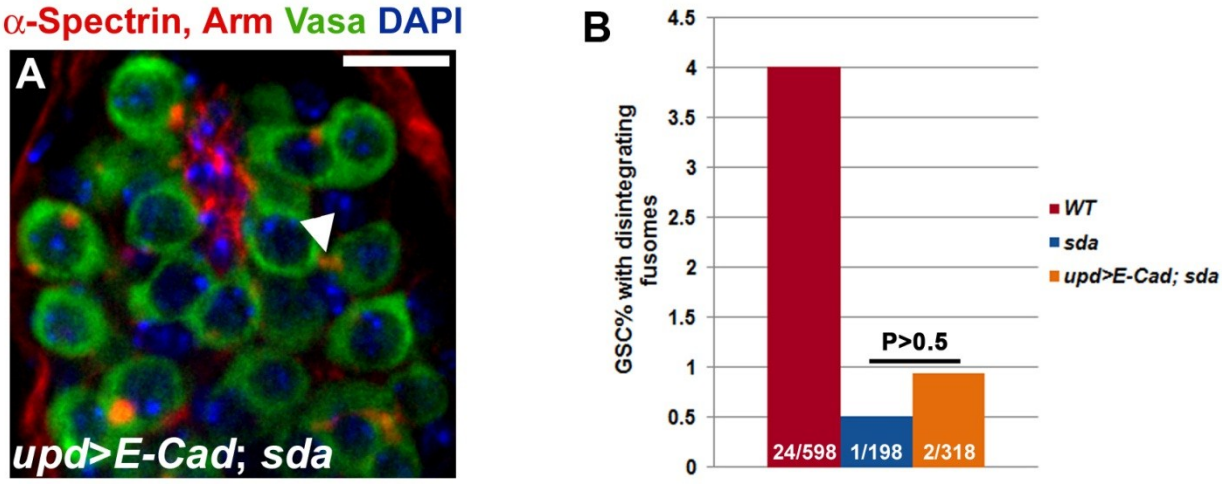


Figure 3-10: Overexpression of *E-Cad* did not rescue dedifferentiation defects in *sda* mutant testes. (**A**) Only the GSC-GB pair connected by a single spectrosome (arrowhead) was detected in testes from *upd-Gal4*; *UAS-E-Cad*; *sda* mutant males (compared to Fig. 3D-E). Scale bar: 10μm. (**B**) Comparison of GSCs with misoriented centrosomes in testes from *sda* mutant males with *upd-Gal4*; *UAS-E-Cad*; *sda* males did not show significant difference (P> 0.5). P-value is calculated by Fisher's Exact Test.

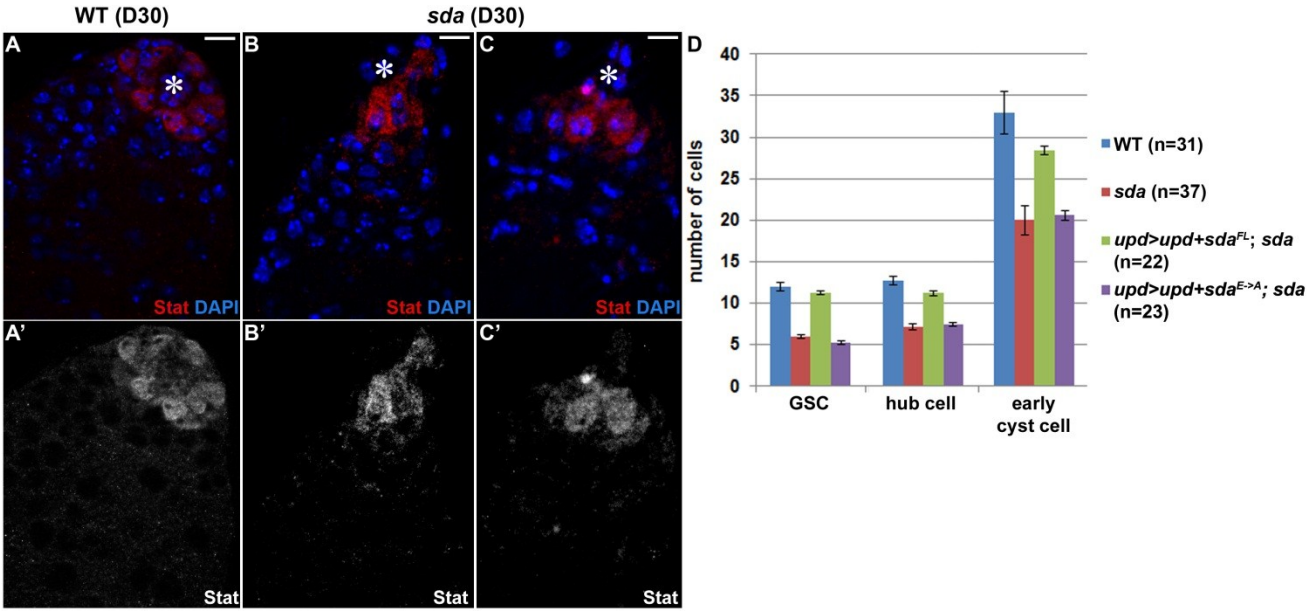


Figure 3-11: E**xpression pattern of Stat in *sda* mutant testes and Upd overexpression in hub cells failed to rescue GSC loss phenotype in *sda* mutant testes.**

Immunostaining using anti-Stat in testes from D30 WT (**A-A'**) and *sda/Df* mutant (**B-B'**, **C-C'**) males. Scale bar: 10μm. (**D**) Quantification of GSC, Zfh-1-positive cell and hub cell number in the testicular niche from WT, or *sda/Df*, or *upd-Gal4*; *UAS-upd*, *UAS-HA-sda*^{FL}; *sda/Df*, or *upd-Gal4*; *UAS-upd*, *UAS-HA- sda*$^{E\rightarrow A}$; *sda/Df* D0-D3 males. For all three cell types, the difference between WT (or *upd-Gal4*; *UAS-upd*, *UAS-HA-sda*^{FL}; *sda/Df)* and *sda/Df* (or *upd-Gal4*; *UAS-upd*, *UAS-HA- sda*$^{E\rightarrow A}$; *sda/Df*) is significant ($P< 0.01$, based on student *t*-test).

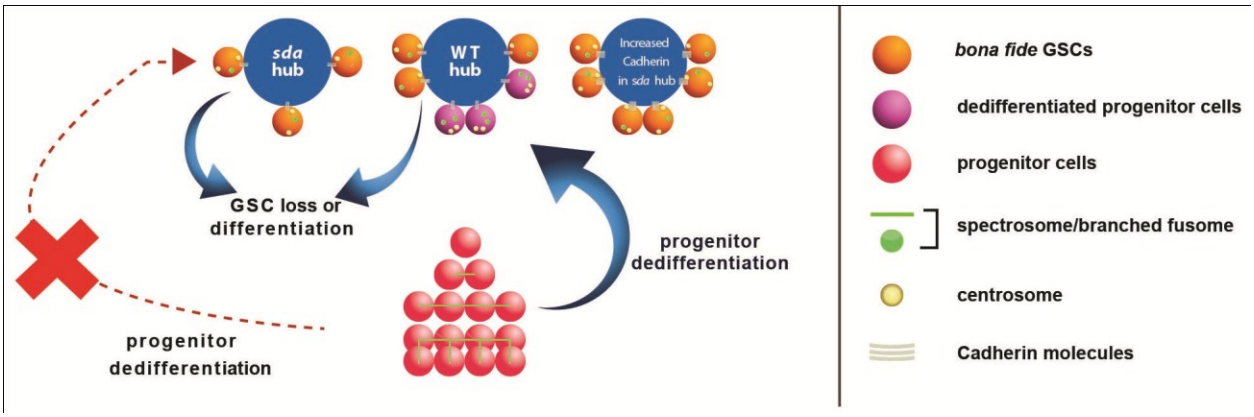


Figure 3-12: A schematic diagram outlines activities of Sda in *Drosophila* testicular niche to maintain GSCs. In WT testicular niche, GSC number is maintained by both self-renewal of *bona fide* GSCs and dedifferentiation of progenitor germ cells, including gonialblasts and spermatogonial cells. In *sda* mutant niche, progenitor cells fail to undergo dedifferentiation, therefore all retained GSCs are *bona fide* GSCs. In *sda* mutant niche with overexpression of Cadherin molecules such as E-Cad and N-Cad, *bona fide* GSCs have increased adhesion to the hub and are lost less frequently, consistent with published work (Boyle et al., 2007; Inaba et al., 2010; Yamashita et al., 2003). Dedifferentiated GSCs can be recognized by misoriented centrosomes and transient disintegrating fusome remnants.

Chapter 4 :

Investigate epigenetic differences between GSCs and its daughter cells during asymmetrical divisions

Summary

Epigenetic mechanisms such as histone modification and the chromatin remodeling have been shown to play important roles in regulating stem cell identity and activity. However, one of the major unanswered questions has been whether and how stem cells retain their epigenetic memory. One of the mechanisms that stem cells use to maintain themselves and to differentiate is to undergo asymmetric cell division. The *Drosophila* germline stem cells (GSC) have been shown to undergo asymmetric cell division. Since the exact mechanisms regulating asymmetrical GSCs division such as physical attachment of stem cells and the mitotic spindle orientation has been very well studied in the testis, it is an excellent system to determine if and how epigenetic memory is retained during asymmetric cell division.

Here we report that stem cells maintain a relatively stable chromatin modification state by retaining pre-existing histones, the key components of chromatin. We found that during the asymmetric division of *Drosophila* male germline stem cell (GSC), the preexisting histones are selectively segregated to the GSC whereas newly synthesized histones during DNA replication are enriched in the differentiating daughter cell. The asymmetric histone inheritance occurs in GSCs but not in symmetrically dividing progenitor cells.

In contrast, histone variant, H3.3 which is incorporated into chromatin in a transcription-dependent manner, partitions symmetrically between GSC and its daughter cell. We also found that if GSCs are forced to divide symmetrically, the asymmetric histone inheritance mode is lost

Thus, our studies demonstrate, for the first time, that canonical histones as the crucial carrier of epigenetic information are asymmetrically inherited in stem cells *in vivo*. Our discoveries significantly advance basic knowledge in the fields of chromatin biology and stem cell biology. These findings provide direct evidence for asymmetric epigenetic inheritance in

stem cells, and offer a critical first step in identifying the cellular machinery involved in this process.

Introduction

Epigenetic mechanisms that alter chromatin structure while preserving primary DNA sequences can regulate gene expression and maintain a specific cell fate through many cell divisions (Jacobs and van Lohuizen, 2002; Ringrose and Paro, 2004; Turner, 2002) such as DNA methylation and chromatin modifying enzymes. However, except for DNA methylation, there is little known about the molecular mechanisms through which these epigenetic changes can be inherited during cell division (Bonasio et al., 2010; Martin and Zhang, 2007).

Stem cells are specialized cells that have the remarkable ability to both self-renew and to generate daughter cells that enter differentiation (Knoblich, 2008). As stem cells supply cells needed for tissue homeostasis during injury and aging process, the maintenance of stem cell identity across many cells divisions is critical. Epigenetic mechanisms have been reported to regulate stem cell activity in multiple lineages, which lead to the hypothesis that stem cells may have a unique chromatin structure (Buszczak and Spradling, 2006; Eun et al., 2010; Jaenisch and Young, 2008). However, to date, there has been little direct *in vivo* evidence which showed whether and how stem cells retain their epigenetic memory.

The *Drosophila* male GSC system is one of the best characterized adult stem cell systems in terms of physiological location, microenvironment (i.e. niche), and cellular structures (Fuller and Spradling, 2007; Losick et al., 2011) (Fig. 4-1A, 4-1B). Male GSCs can be identified precisely by their distinct anatomical positions and morphological features. A GSC usually divides asymmetrically to give rise to a self-renewed GSC and GB that undergoes differentiation. Therefore, GSCs can be examined at single-cell resolution. This resolution will allow for a direct comparison between the two daughter cells from GSC division.

Since DNA methylase activity is almost negligible in adult flies (Hung et al., 1999; Lyko et al., 2000a; Lyko et al., 2000b; Richards and Elgin, 2002), we postulated that histones are one

of the major carriers of epigenetic information (Kouzarides, 2007) in this system. Thus, we focused on experiments that address how histones are inherited during the asymmetric division of GSC.

Materials and Methods

Fly strains and husbandry

Fly stocks were raised using standard Bloomington medium at 18°C and heat shocked for two hours at 37°C in a circulating water bath. The following fly stocks were used: *hs-FLP* on the X chromosome (Bloomington Stock Center BL-26902), *nos-gal4* on the 2nd chromosome (Van Doren et al., 1998), and *UAS-upd* on the 2nd chromosome (Terry et al., 2006).

Generation of switchable dual-color transgenic fly strains

Standard procedures were used for all subcloning experiments. Enzymes used for plasmid construction were obtained from New England Biolabs (Beverly, MA) and Promega Biotech (Madison, WI). *H3* sequences were excised from the *K90* plasmid (Ahmad and Henikoff, 2002a) as an *Xba*I and *Eag*I flanked fragment, and *mKO* sequences were recovered using PCR from a *UAS-mKO-vasa* transgenic fly strain (a gift from Akira Nakamura) as an *Xba*I and *Eag*I flanked fragment. *H3* and *mKO* fragments were then inserted by a three-way ligation into the *Xba*I site of the *UASp* plasmid to construct the *UASp-H3-mKO* plasmid (Rorth, 1998). *H3-GFP* fusion sequences were retrieved from the *K90* plasmid (Ahmad and Henikoff, 2002a) as an *Xba*I-*Dra*I fragment followed by treatment with Klenow (NEB Cat# M0210), which was subsequently subcloned into an *Nhe*I site (treated with Klenow) between two *FRT* sites and upstream of the *SV40 PolyA* sequences. The entire *FRT-H3-GFP-SV40 PolyA-FRT* sequences were then

recovered using *Not*I and *Bgl*II, followed by subcloning into the *UASp-H3-mKO* plasmid, as described previously, digested by *Not*I and *BamH*I (note: *Bgl*II and *BamH*I produce compatible cohesive ends). Similar cloning strategies were used for *H2B* and *H3.3*, except that the *H2B* sequences were recovered from the *K161* plasmid and the *H3.3* sequences were recovered from the *K48* plasmid. All histone plasmids were generously provided by Kami Ahmad (Harvard University). The final *UASp* plasmids were introduced to w^{1118} flies by P-element-mediated germline transformation (Bestgene Inc.).

Immunofluorescence

Immunofluorescence staining was performed as described previously (Hime et al., 1996). Primary antibodies were anti-α spectrin (1:20, DHSB) and anti-Fasciclin III (1:10, DHSB). Images were taken using the Zeiss LSM 510 META or Zeiss LSM 510 Multiphoton confocal microscope and processed using NIH Image J and Adobe Photoshop software.

Quantification of images

No antibody was added to enhance either GFP or mKO signal. Values of GFP and mKO intensity were calculated using Image J software: DAPI signal was used to determine the area of nucleus for measuring both GFP and mKO fluorescent signals, the raw reading was subsequently adjusted by subtracting background fluorescence signals in both GSC and GB nuclei and compared between each other (Table S1-S5, Fig. 2H and Fig. 3F).

Results

Histones are one of the major carriers of epigenetic information (Kouzarides, 2007). To address how histones are inherited during the asymmetric division of GSC, we developed a novel switchable dual-color method to differentially label "old" vs. "new" histones (Fig. 4-1C). This method employs both spatial (by Gal4; UAS system) and temporal (by heat shock induction) controls to switch labeled histones from green [Green Fluorescent Protein (GFP)] to red [monomeric Kusabira-Orange (mKO)]. Heat shock treatment induces an irreversible DNA recombination to shut down expression of the old histones labeled with GFP and initiate expression of the new histones labeled with mKO.

We postulated that there are two possible outcomes: 1. If the old histones are inherited non-selectively, the green fluorescent histones will initially exhibit equal distribution in the self-renewed GSC and the differentiating GB, and will be gradually replaced by the red fluorescent histones (Fig. 4-1D). 2. If the old histones are preferentially retained in the GSCs to constitute potentially GSC-specific chromatin structure, the green fluorescent histones would be retained specifically in the GSCs (Fig. 4-1E).

During DNA replication-dependent histone deposition, the histones H3 and H4 are incorporated as a tetramer, and histones H2A and H2B are incorporated as dimers (Annunziato et al., 1982; Jackson and Chalkley, 1981a; Jackson and Chalkley, 1981b; Russev and Hancock, 1981; Xu et al., 2010). Thus, we generated independent transgenic fly strains for the histones H3 and H2B. On the other hand, histone variants are incorporated into chromatin in a transcription-coupled, but DNA replication-independent manner (Ahmad and Henikoff, 2002b; Tagami et al., 2004). Therefore, the histone variant H3.3 was also used to generate transgenic strains as a control for canonical histones.

To avoid potential complications caused by heat shock-induced DNA recombination on either one or both chromosomes in GSCs, each of the three transgenes (H3, H2B and H3.3) was integrated as a single copy and analyzed in heterozygous flies. Examination of testes from transgenic males revealed nuclear GFP but no mKO signal before heat shock(Figure 4-2). After the heat shock treatment, mKO signals were detectable in germ cell nuclei (Fig. 4-2).

Different GSCs undergo mitosis asynchronously, and an average cell cycle length of GSCs is approximately 12 to 16 hours. Among all GSCs, 75-77% are in G2 phase, 21% are in S phase, less than 2% are in mitosis, and GSCs in G1 phase are almost negligible (Cheng et al., 2008; Sheng and Matunis, 2011; Yadlapalli et al., 2011; Yamashita et al., 2003; Yamashita et al., 2007). Moreover, the GSC and GB arising from an asymmetric division remain connected after mitosis by a cellular structure known as the spectrosome, when they undergo the next G1 and S phases synchronously (Sheng and Matunis, 2011; Yadlapalli et al., 2011).

To compare the distribution of old and new histones in mitosis after one round of DNA replication-dependent histone deposition, testes were studied 16 to 20 hours after heat shock (Fig. 4-3A). In particular, GSC-GB pairs that were connected by spectrosomes were examined (arrows in Fig. 4-3B, 4-3D). Based on the cell cycle length of GSCs, these GSC-GB pairs were derived from GSCs that switched from *histone-GFP* to *histone-mKO* genetic code by heat shock treatment during G2 phase, underwent the first mitosis followed by G1, S, G2 phase and the second mitosis (Fig. 4-3A). With this time frame, both old histones and new histones were detectable in the G2 phase GSCs (Fig. 4-3E-E", Table 4-2), because new histones have been synthesized and incorporated into chromatin after the S phase. For the canonical histone H3, the GFP signal was detected primarily in the GSC, but not in the GB (Fig. 4-3B'). By contrast, the mKO signals were present in both the GSC and the GB, with a relatively higher level in the GB (Fig. 4-3B, 4-3B"). The asymmetric inheritance mode of histone H3 was specific for GSC divisions, because both the GFP and the mKO signals were equally distributed in spermatogonial

cells derived from a symmetric division of the GB, in the same testis sample (Fig. 4-3C-C"). Quantification of fluorescence intensity revealed that the old H3 (GFP-labeled) signal was ~5.5-fold more enriched in the GSC compared to the GB, while new H3 (mKO-labeled) signal was ~1.5-fold more enriched in the GB compared to the GSC (H3 GSC/GB data in Fig.4-3H, Table 4-1 and 4-2). Consistent with the fluorescence image data, this differential distribution of old vs. new histone was not detected for symmetrically dividing spermatogonial cells (H3 SG1/SG2 data in Fig.4-3H, Table S4-1 and 4-2: ratio of H3-GFP in SG1/SG2= 1.09; ratio of H3-mKO in SG1/SG2= 1.02).

In contrast to the asymmetric inheritance pattern for the canonical histone H3, the histone variant H3.3 showed symmetric pattern during GSC divisions, by fluorescence images (Fig. 4-3D-D") and by quantification of fluorescence intensity in GSC-GB pairs expressing the H3.3 transgene (H3.3 GSC/GB data in Fig. 2H: ratio of H3.3-GFP in GSC/GB= 1.03; ratio of H3.3-mKO in GSC/GB= 1.03, Table S4-1 and 4-2). The symmetric distribution pattern of the histone variant H3.3 suggested that the asymmetric inheritance mode is specific for canonical histones.

Because mitotic GSCs are less than 2% of the total scored GSCs, all analysis above were based on post-mitotic GSC-GB pair. To further examine whether asymmetric histone inheritance is through asymmetric histone segregation during mitosis, we screened for mitotic GSCs at both metaphase and anaphase/telophase. Strikingly, association of old histones to the chromatids that will be segregated to the GSC was already detectable at metaphase (Fig.4-3F-F"), and became distinct at anaphase (Fig.4-3G-G'). By contrast, new histones were more enriched at the chromatids that are segregated to GB (Fig.4-3G, 4-3G"). These results suggest that the sister chromatids preloaded with old histones are preferentially retained in GSCs and the sister chromatids enriched with new histones partitioned to GBs during GSC asymmetric division.

Next, to examine the histone inheritance mode during the first GSC division, GSCs were recovered for 4 to 6 hours following heat shock (Fig. 4-4A). Interestingly, an asymmetric inheritance mode was found in GSC-GB pairs with the H3 transgene (Fig.4-4B-B"). By contrast, symmetric inheritance mode was observed for both dividing spermatogonial cells with the H3 transgene (Fig. 4-4C-C") and H3.3 during GSC division (Fig. 4-4D-D"). Quantification of fluorescence intensity revealed that the old H3-GFP signal was ~11.7-fold more enriched in the GSC compared to the GB, while the new H3-mKO signal was ~2.4-fold more enriched in the GB compared to the GSC (H3 GSC/GB data in Fig.4-4F, Table 4-3 and 4-4). By contrast, there was no differential distribution of the old vs. new histone for the symmetrically dividing spermatogonial cells (H3 SG1/SG2 data in Fig.4-4F, Table 4-3-4-4: ratio of H3-GFP in SG1/SG2=1.07; ratio of H3-mKO in SG1/SG2=1.06), or H3.3 during GSC division (H3.3 GSC/GB data in Fig. 4-4F, Table 4-3-S4-4: ratio of H3.3-GFP in GSC/GB =1.00; ratio of H3.3-mKO in GSC/GB =1.02). Although asymmetric distribution of old vs. new histones was detected in post-mitotic GSC-GB pair, examination of the mitotic GSC at this stage did not show any asymmetric distribution (Fig.4-4E-E"'), suggesting that the asymmetric histone segregation mode (Fig. 4-3G-G") relies on replication-dependent histone incorporation prior to mitosis. However, this more than 10-fold difference of GFP signal in GSC and GB could be explained by rapid turn-over of old histones in GBs. By contrast, the difference of mKO in GSC and GB was less dramatic, probably due to newly synthesized histones in both cells. Furthermore, results from the H2B transgenic males were similar to those of H3 (fig. 4-5), suggesting a common mode of inheritance for canonical histones. Lastly, we postulated that GB prefer newly synthesized histone as a mechanism to reset the chromatin for differentiation. This hypothesis is also supported by our finding that GB is enriched in chromatin factors transcripts compared to the samples that contain stem cells (Chapter 2).

Although male GSCs normally undergo asymmetric cell divisions to self-renew and generate GBs that undergo differentiation, this asymmetry is lost under certain conditions, such as ectopic activation of the key JAK-STAT (Janus kinase and signal transducer and activator of transcription) signaling pathway in the niche (Kiger et al., 2001; Leatherman and Dinardo, 2008; Tulina and Matunis, 2001). It has been shown that overexpression of the JAK-STAT ligand *unpaired* (OE-*upd*) induces overpopulation of GSCs (Kiger et al., 2001; Tulina and Matunis, 2001). Consistent with the loss of asymmetry in expanded GSCs, the asymmetric inheritance mode of the histone H3 was not observed in dividing GSCs from OE-*upd* testes 16-20 hour after heat shock (Fig.4-6A-A"). These results demonstrate that the asymmetric histone inheritance pattern is dependent on an asymmetric mode of GSC division, which produces two daughter cells with distinct cell fates. We summarize our data in the context of GSC cell cycle (Fig.4-6B, an alternative situation discussed in fig. 4-7) and propose that the asymmetric histone inheritance occurs in two steps: old and newly synthesized histones are incorporated to different sister chromatids during S phase, then during mitosis, the sister chromatid with old histones is preferentially segregated to GSC.

Discussion

Understanding the molecular mechanisms underlying the ability of stem cells to maintain their unique features is vital to exploit their roles in tissue homeostasis and therapeutic potential in tissue regeneration. Our work reveals that stem cells preserve preexisting histones through asymmetric cell divisions. Retention of preexisting histone in stem cells suggests that these histones may play a role in maintaining its stem cell function. The mechanism of how these histones are retained remains to be determined. Moreover, it is unclear whether these histones are retained due to their post-translational modifications functioning as a "histone code". We

hypothesized that the histones modification in these preexisting histones need to be passed on to the newly synthesized histone. However, the perdurance of this process throughout the cells' lifetime still need to be determined.

Our results also show that the JAK-STAT signaling pathway is required for the asymmetric mode of stem cell divisions and contributes to the asymmetric inheritance pattern of histones. As JAK-STAT is a major signaling pathway in this stem cell system, there still remains questions of how this signaling pathway cross-talks with the mechanism that retains the preexisting histones. However, we conclude that this work provides a critical first step toward identifying the detailed molecular mechanisms responsible for retaining old histones during GSC asymmetric divisions. Lastly, this finding in the well characterized GSC model system can be generally applied to other stem cell systems to understand how epigenetic information can be maintained by stem cells, or reset in their sibling cells which undergo cellular differentiation.

Figures and Tables

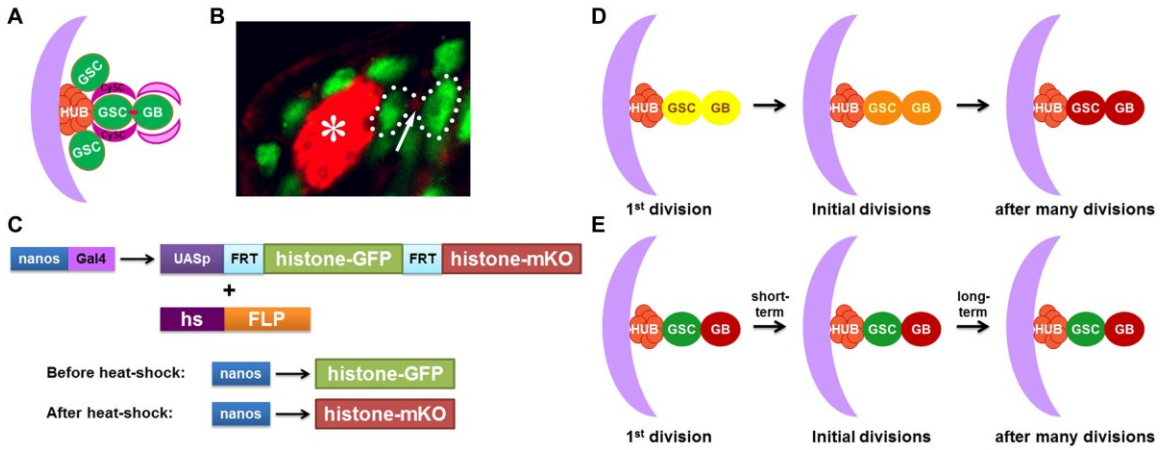


Figure 4-1: Experimental design and potential results of histone inheritance pattern during GSC asymmetric cell division. (**A**) A cartoon image of the *Drosophila* male germline stem cell niche at the tip of testis. HUB- hub cells, GSC- germline stem cell, CySC- cyst progenitor or somatic stem cell, GB- gonialblast. (**B**) Immunofluorescence image showing the structure of GSC niche: HUB (anti-Fas III, red, asterisk), GSC-GB pairs expressing H3-GFP (green) connected by spectrosome (anti- Immunofluorescence imag**(C)** Structure of the transgene *UASp-FRT-histone-GFP-PolyA-FRT-histone-mKO-PolyA*. UAS: upstream activating sequence; FRT: FLP (flipase) recombination target; histone: H3, H2B, or histone variant H3.3. *nanos-Gal4*: a germline-specific Gal4 driver. *hs-FLP*: the yeast FLP recombinase is under the control of a heat shock promoter (*hs*). **(D-E)** Predicted results of histone segregation during GSC division: for simplicity, only one GSC-GB pair is shown, and each entire cell is colored to represent histone fluorescent signals in cartoon.

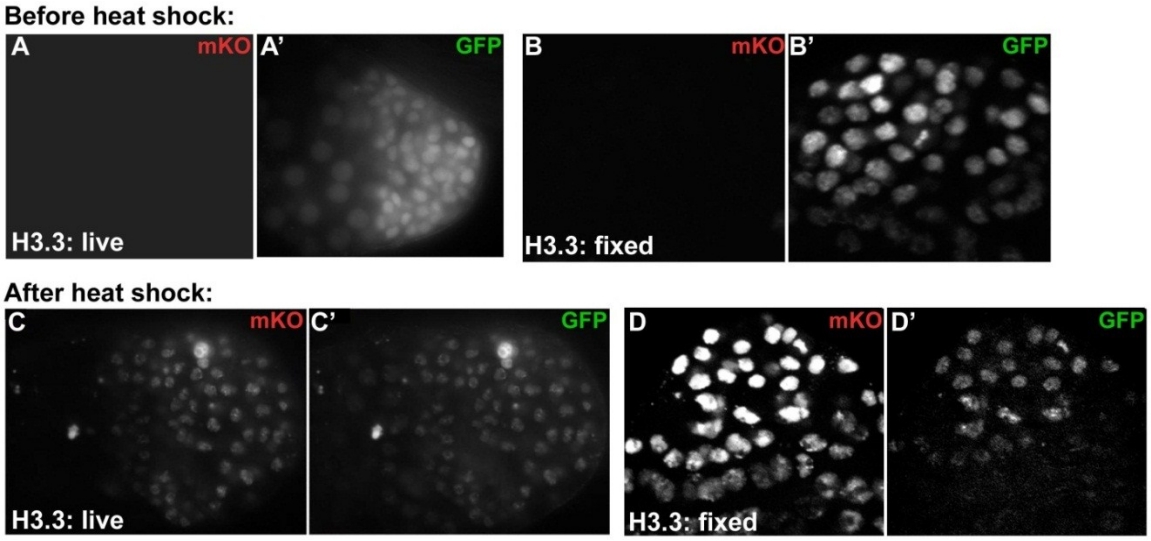


Figure 4-2: GFP and mKO signals before and after heat shock induction in H3.3 transgenic male testis samples. (A-A') In live samples before heat shock, mKO was undetectable (**A**) and GFP was detectable in early germ cells (**A'**). **(B-B')** In fixed samples before heat shock, mKO was undetectable (**B**) and GFP was detectable in early germ cells (**B'**). (**C-C'**) In live samples after heat shock, both mKO (**C**) and GFP (**C'**) signals were detectable. (**D-D'**) In fixed samples after heat shock, both mKO (**D**) and GFP (**D'**) signals were detectable. All images shown in Figure 2-4 were fixed samples after heat shock. Therefore fixation did not change the GFP or mKO pattern.

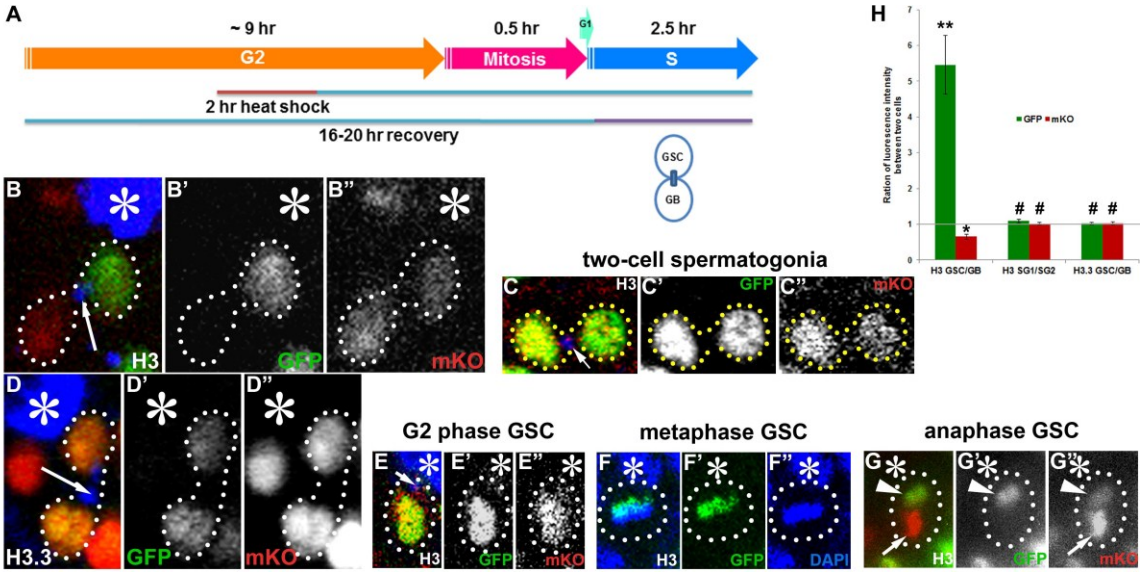


Figure 4-3: Canonical histone H3 is asymmetrically segregated during the second GSC division after heat shock. (A) Heat shock regime. **(B-B")** H3 is inherited asymmetrically in GSC vs. GB. **(C-C")** H3 is inherited symmetrically in two-cell spermatogonia. (**D-D"**) H3.3 is inherited symmetrically in GSC vs. GB. (**E-E"**) Both H3-GFP and H3-mKO were detectable in GSCs at G2 phase. (**F-F"**) Metaphase GSC: H3-GFP was localized to the chromatids that were segregated to GSCs. (**G-G"**) Anaphase GSC: H3-GFP was associated with the condensed chromosomes that were segregated into GSC (arrowhead), while H3-mKO was enriched at the condensed chromosomes going to GB (arrow). Asterisk: HUB; arrow: spectrosome. (**H**) Quantification of GFP and mKO fluorescence intensity ratio (Table 4-2). H3 GSC-GB pair (N=14): GSC/GB GFP ratio is significantly > 1 (** $P<10^{-4}$), GSC/GB mKO ratio is significantly < 1 (* P=0.0002). H3 two-cell spermatogonial (SG) pair (N=16): neither SG1/SG2 GFP (# P=0.103) nor mKO (# P=0.684) ratio is significantly different from 1. H3.3 GSC-GB pair (N=12): neither GSC/GB GFP (# P=0.513) nor mKO (# P=0.532) ratio is significantly different from 1. All data were 16-20 hour after heat shock (raw data in Table 4-1). Error bars: standard error. P-value is calculated by one sample t-test.

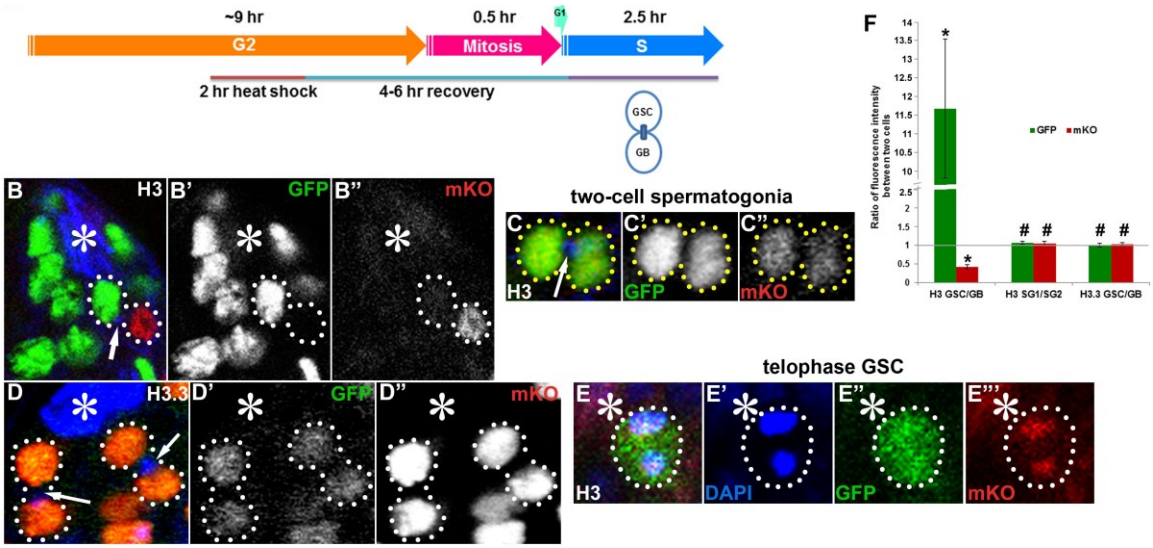


Figure 4-4: Canonical histones H3 is asymmetrically inherited during the first GSC division after heat shock. (A) Heat shock regime. **(B-B")** H3 is inherited asymmetrically in GSC vs. GB. **(C-C")** H3 is inherited symmetrically in two-cell spermatogonia. (**D-D"**) H3.3 is inherited symmetrically in GSC vs. GB. Asterisk: HUB; arrow: spectrosome. (**E-E"'**) Telophase GSC: both GFP and mKO were symmetrically segregated. (**F**) Quantification of GFP and mKO fluorescence intensity ratio (Table 4-4). H3 GSC-GB pair (N=13): GSC/GB GFP ratio is significantly > 1 (* $P< 10^{-4}$), GSC/GB mKO ratio is significantly < 1 (* $P< 10^{-4}$). H3 two-cell spermatogonial (SG) pair (N=11): neither SG1/SG2 GFP (# P=0.225) nor mKO (# P=0.365) ratio is significantly different from 1. H3.3 GSC-GB pair (N=13): neither GSC/GB GFP (# P=0.970) nor mKO (# P=0.594) ratio is significantly different from 1. All data were obtained at 4-6 hours after heat shock (raw data in Table 4-3). Error bars: standard error. P-value is calculated by one sample t-test.

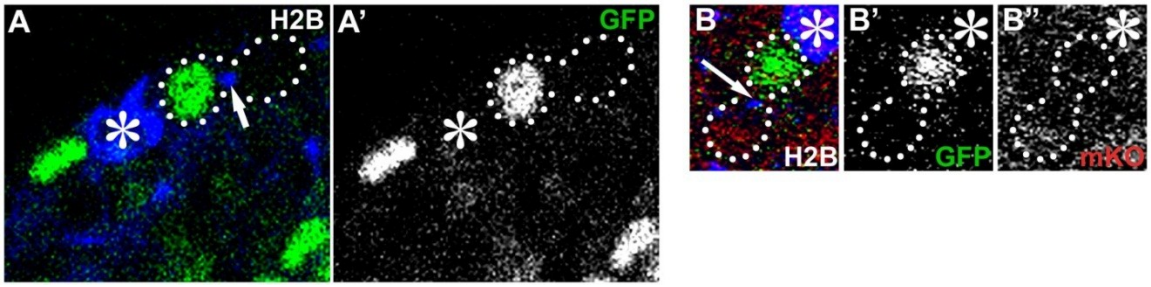


Figure 4-5 : H2B inheritance pattern very much resembled the H3 pattern. (A-A') (B-B")

H2B is inherited asymmetrically in GSC vs. GB. GFP signal was only detectable in GSC but not in GB **(A', B')**, whereas mKO signal is more enriched in GB than in GSC **(B")**.

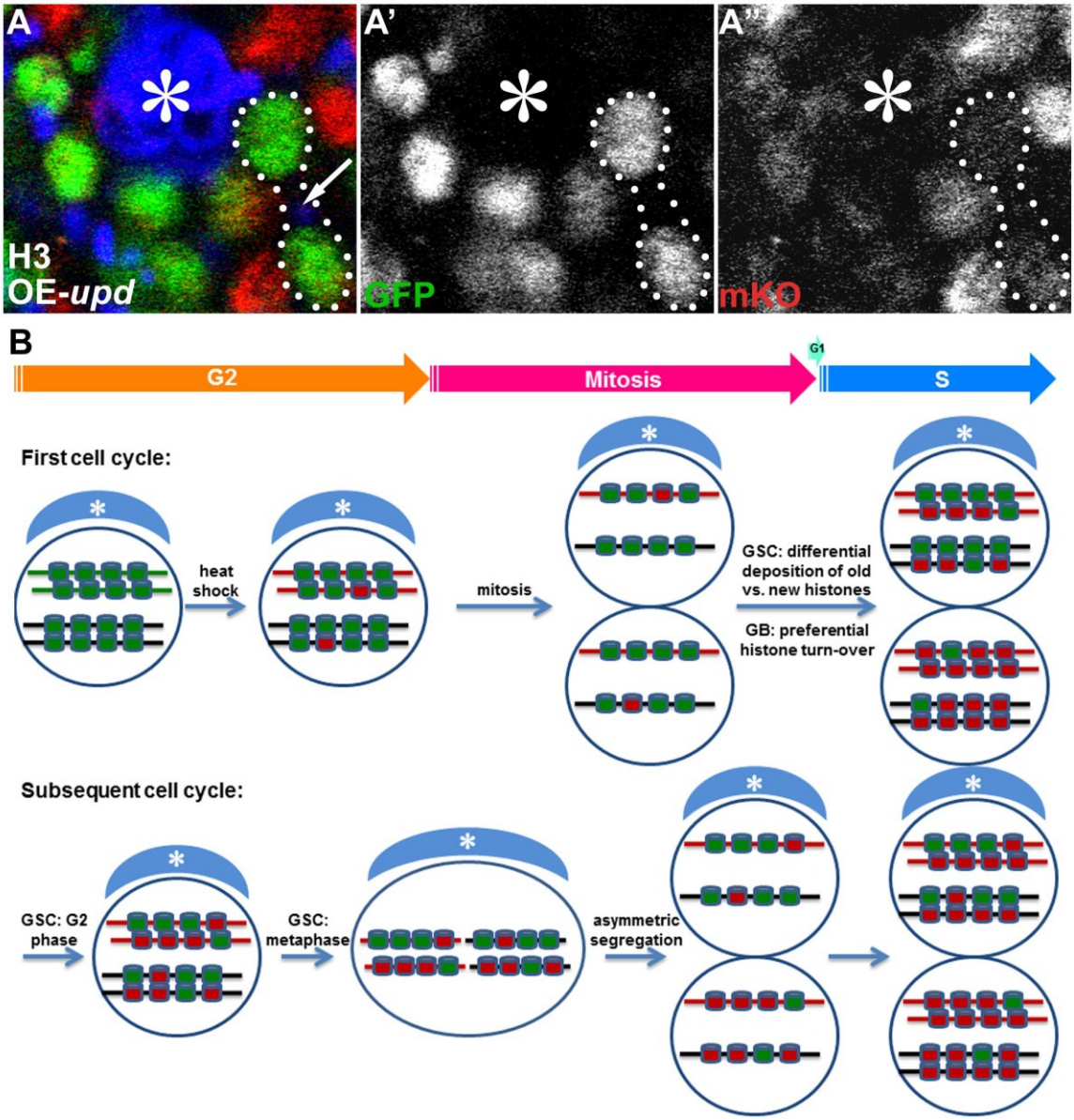


Figure 4-6: The asymmetric histone inheritance mode is abolished in *upd* overexpression testes and interpretation of the data. (A-A") Symmetric distribution of old H3-GFP (**A'**) and new H3-mKO (**A"**) in the two daughter cells from a GSC division in *nanos-Gal4*; *UAS-upd* males. **(B)** Interpretation of the data in the context of GSC cell cycle: each line represents one sister chromatid. Black line is the chromosome without transgene; green line is the chromosome with the transgene (Fig.4-1C) before heat-shock induced DNA recombination; red line is the

chromosome with the transgene after heat-shock induced DNA recombination. The green cylinder represents nucleosome containing old histone-GFP, and red cylinder is for nucleosome with new histone-mKO (only canonical histones are discussed here). Some old histones could be replaced by new histones during G2 phase (Dion et al., 2007) but not at the global level. For simplicity, nucleosomes with both histone-GFP and histone-mKO, or histones that have not been incorporated into chromatin yet, are not discussed here.

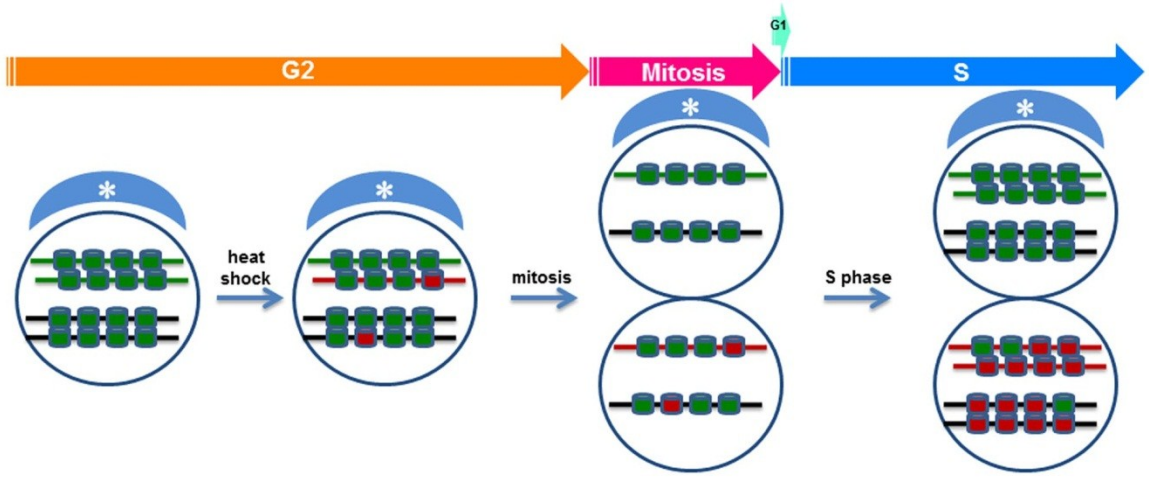


Figure 4-7: Alternative interpretation of the data in the context of GSC cell cycle. If only one sister chromatid carrying the transgene was flipped and changed from *histone-GFP* to *histone-mKO* genetic code in G2 phase GSCs. This scenario is unlikely because the same chance exists for the opposite chromatid segregation mode, i.e., chromatid with *histone-mKO* goes to GSC, while chromatid with *histone-GFP* goes to GB, which should produce mKO mainly in GSC, and GFP mainly in GB. However, this phenomenon has never been observed. In addition, this scenario will lead to GSCs with the *histone-GFP* transgene, which should only give rise to GFP-labeled but not mKO-labeled histones. But GSCs with both GFP-labeled and mKO-labeled histones at G2 phase in the following cell cycle were observed (Fig. 2E-E").

Table 4-1: Raw data in 16-20hr chasing experiments after heat shock treatment.


#	H3 GSC-GB pair		H3 two-cell gonial		H3 G2 phase		H3.3 GSC-GB pair	
1	GSC	1500.099	Cell 1	1397.613	GSC	1401.806	GSC	285.215
		291.183		330.821				1014.544
	GB	142.511	Cell 2	1387.796		454.398	GB	264.673
		466.943		301.563				825.591
2	GSC	1529.681	Cell 1	1418.267	GSC	1265.373	GSC	196.707
		136.623		417.317				676.726
	GB	319.211	Cell 2	1355.055		409.654	GB	260.726
		585.575		446.387				918.63
3	GSC	435.249	Cell 1	188.7	GSC	517.538	GSC	283.127
		287.555		40.685				770.667
	GB	129.322	Cell 2	236.416		288.972	GB	291.394
		386.614		51.116				724.056
4	GSC	865.323	Cell 1	1151.203	GSC	393.291	GSC	255.564
		365.887		728.587				740.029
	GB	111.639	Cell 2	917.165		714.193	GB	250.216
		851.825		614.398				745.186
5	GSC	942.64	Cell 1	880.966	GSC	1152.725	GSC	2062.419
		535.296		861.091				1289.867
	GB	140.525	Cell 2	595.153		671.884	GB	2026.555
		913.644		733.017				1190.467
6	GSC	928.09	Cell 1	664.26	GSC	1262.267	GSC	912.084
		416.658		747.059				3428.707
	GB	168.942	Cell 2	622.782		679.617	GB	782.29
		897.531		657.118				3184.586
7	GSC	312.252	Cell 1	480.14	GSC	535.575	GSC	952.056
		248.17		623.93				3772.343
	GB	28.58	Cell 2	360.393		93.178	GB	1019.955
		323.363		566.118				3755.15
8	GSC	121.038	Cell 1	157.33	GSC	276.904	GSC	3128.038
		145.349		171.068				2667.677
	GB	23.446	Cell 2	135.381		218.829	GB	3258.834
		129.898		150.069				2810.174

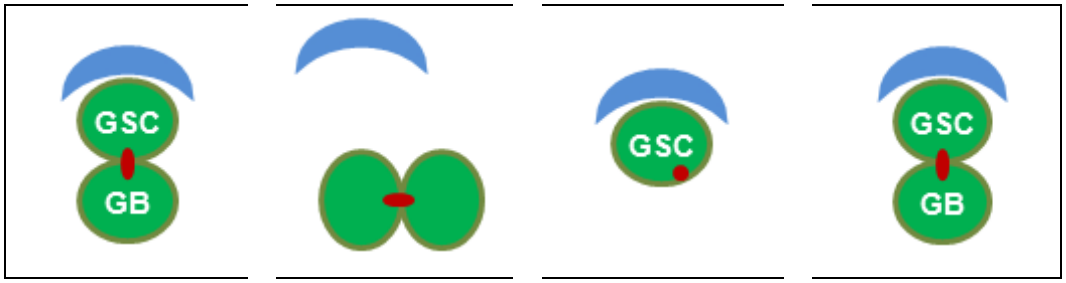
	GSC	158.333		516.145				1040.175
		119.02	Cell 1	359.614		39.134	GSC	3812.473
		48.848		335.694				956.841
9	GB	141.182	Cell 2	280.728	GSC	250.698	GB	3833.674
	GSC	180.352		437.637				893.77
		192.303	Cell 1	305.594		649.452	GSC	3391.099
		81.708		396.481				962.749
10	GB	891.491	Cell 2	340.233	GSC	200.867	GB	3912.301
	GSC	1517.473		729.995				946.844
		703.033	Cell 1	364.953		206.963	GSC	3703.863
		177.875		648.624				798.391
11	GB	842.059	Cell 2	366.977	GSC	169.836	GB	3673.802
	GSC	365.388		990.755				1013.492
		376.23	Cell 1	281.937		238.279	GSC	3214.758
		230.469		1171.972				846.325
12	GB	739.583	Cell 2	349.169	GSC	195.984	GB	2395.879
	GSC	621.364		213.948				
		421.954	Cell 1	188.766		1103.981		
		141.598		242.074				
13	GB	519.144	Cell 2	230.611	GSC	686.025		
	GSC	405.951		520.058				
		379.202	Cell 1	404.64		1021.287		
		239.143		560.002				
14	GB	377.236	Cell 2	416.185	GSC	372.79		
				1189.539				
			Cell 1	541.433		287.769		
				1193.312				
15			Cell 2	505.559	GSC	582.587		
				1001.169				
			Cell 1	390.725		196.403		
				1078.839				
16			Cell 2	465.487	GSC	82.243		

Table 4-1: Fluorescence intensity of H3 GFP and mKO in GSC vs. GB, H3 GFP and mKO in spermatogonia (SG) 1 vs. SG 2, H3 GFP and mKO in G2 phase GSCs, H3.3 GFP and mKO in GSC vs. GB, 16-20 hour post heat shock treatment. Green highlighted data are for GFP signal, pink highlighted data are for mKO signal. No antibody was used to enhance fluorescence signals.

Table S2: Summary of the data in 16-20hr chasing experiments after heat shock for Figure 2H.						
	H3 GSC/GB	H3 SG1/SG2	H3.3 GSC/GB	S.E.(H3 GSC/GB)	S.E.(H3 SG1/SG2)	S.E.(H3.3 GSC/GB)
GFP	5.455382459	1.093612642	1.025073891	0.820475593	0.054112915	0.036950228
mKO	0.656220573	1.015601146	1.029019606	0.071721656	0.038523513	0.044883253

Table 4-2: Summary of the data in longer-term chasing experiments

Table 4-3: Raw data in 4-6hr chasing experiments after heat shock treatment.



#	H3 GSC-GB pair		H3 two-cell gonial		H3 G2 phase		H3.3 GSC-GB pair	
		2224.035		619.872				97.941
	GSC	158.255	Cell 1	771.852		1739.536	GSC	2206.255
		130.344		458.859				112.722
1	GB	416.013	Cell 2	734.225	GSC	99.933	GB	2578.891
		2483.218		123.883				815.022
	GSC	359.261	Cell 1	3012.525		2880.902	GSC	3124.291
		123.707		96.526				960.713
2	GB	1538.096	Cell 2	2651.048	GSC	327.19	GB	3780.433
		61.378		420.982				976.13
	GSC	12.819	Cell 1	861.173		632.338	GSC	3793.243
		2.588		559.676				826.487
3	GB	30.512	Cell 2	870.169	GSC	435.945	GB	3453.721
		1507.757		1111.096				936.795
	GSC	272.196	Cell 1	788.81		1928.43	GSC	3913.044
		98.831		928.897				1133.497
4	GB	843.043	Cell 2	810.529	GSC	245.314	GB	4038.758
		290.853		3524.34				924.133
5	GSC	104.439	Cell 1	698.885	GSC	955.023	GSC	3698.133

		55.963		3317.38				950.278
	GB	131.534	Cell 2	457.829		79.693	GB	3929.375
		524.237		2035.577				1931.467
	GSC	64.787	Cell 1	88.119		1071.436	GSC	1247.392
		80.529		1766.823				1896.342
6	GB	225.956	Cell 2	68.511	GSC	69.562	GB	1186.402
		701.129		2448.66				3089.152
	GSC	82.682	Cell 1	237.496		789.924	GSC	2589.167
		86.618		2469.899				3337.602
7	GB	159.093	Cell 2	234.81	GSC	68.229	GB	2877.842
		932.499		2727.286				3483.646
	GSC	107.458	Cell 1	202.021		729.871	GSC	2161.199
		69.637		2817.974				3389.671
8	GB	133.159	Cell 2	210.216	GSC	95.088	GB	2057.182
		2725.749		3007.404				893.532
	GSC	651.351	Cell 1	323.327		2018.215	GSC	3390.044
		791.951		2940.333				920.61
9	GB	1192.171	Cell 2	349.009	GSC	638.479	GB	3659.989
		2191.801		3535.573				1012.585
	GSC	513.333	Cell 1	512.304		2145.188	GSC	4068.886
		565.407		3586.377				1175.963
10	GB	846.212	Cell 2	588.952	GSC	749.946	GB	4064.611
				2781.275				2547.362
			Cell 1	243.681		2287.577	GSC	3051.526
				2906.644				3224.452
11			Cell 2	274	GSC	102.398	GB	2818.903
								639.407
							GSC	3329.088
								449.305
12							GB	2341.38
								631.352
							GSC	3288.112
								502.314
13							GB	2757.68

Table 4-3: Fluorescence intensity of H3 GFP and mKO in GSC vs. GB, H3 GFP and mKO in spermatogonia (SG) 1 vs. SG 2, H3 GFP and mKO in G2 phase GSCs, H3.3 GFP and mKO in GSC vs. GB, 4-6 hour post heat shock treatment. Green highlighted data are for GFP

signal, pink highlighted data are for mKO signal. No antibody was used to enhance fluorescence signals.

Table 4-4: Summary of the data in 4-6hr chasing experiments after heat shock for Figure 3F.						
	H3 GSC/GB	H3 SG1/SG2	H3.3 GSC/GB	S.E.(H3 GSC/GB)	S.E.(H3 SG1/SG2)	S.E.(H3.3 GSC/GB)
GFP	11.67681652	1.065625336	0.997746128	1.863207233	0.05105125	0.05186895
mKO	0.424737748	1.056478261	1.024303666	0.061136912	0.058955835	0.043844591

Table 4-4: Summary of the data in short-term chasing experiments

References

1.Ahmad, K. and Henikoff, S. (2002a). Histone H3 variants specify modes of chromatin assembly. *Proc Natl Acad Sci U S A* 99 Suppl 4, 16477-84.

2.Ahmad, K. and Henikoff, S. (2002b). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* 9, 1191-200.

3.Ai, X. and Parthun, M. R. (2004). The nuclear Hat1p/Hat2p complex: a molecular link between type B histone acetyltransferases and chromatin assembly. *Mol Cell* 14, 195-205.

4.Annunziato, A. T., Schindler, R. K., Riggs, M. G. and Seale, R. L. (1982). Association of newly synthesized histones with replicating and nonreplicating regions of chromatin. *J Biol Chem* 257, 8507-15.

5.Ayyar, S., Jiang, J., Collu, A., White-Cooper, H. and White, R. A. (2003). Drosophila TGIF is essential for developmentally regulated transcription in spermatogenesis. *Development* 130, 2841-52.

6.Barreau, C., Benson, E., Gudmannsdottir, E., Newton, F. and White-Cooper, H. (2008a). Post-meiotic transcription in Drosophila testes. *Development* 135, 1897-902.

7.Barreau, C., Benson, E. and White-Cooper, H. (2008b). Comet and cup genes in Drosophila spermatogenesis: the first demonstration of post-meiotic transcription. *Biochem Soc Trans* 36, 540-2.

8.Barroca, V., Lassalle, B., Coureuil, M., Louis, J. P., Le Page, F., Testart, J., Allemand, I., Riou, L. and Fouchet, P. (2009). Mouse differentiating spermatogonia can generate germinal stem cells in vivo. *Nat Cell Biol* 11, 190-6.

9.Beall, E. L., Lewis, P. W., Bell, M., Rocha, M., Jones, D. L. and Botchan, M. R. (2007). Discovery of tMAC: a Drosophila testis-specific meiotic arrest complex paralogous to Myb-Muv B. *Genes Dev* 21, 904-19.

10.Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature* 447, 407-12.

11.Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K. et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315-26.

12.Betschinger, J. and Knoblich, J. A. (2004). Dare to be different: asymmetric cell division in Drosophila, C. elegans and vertebrates. *Curr Biol* 14, R674-85.

13.Bonasio, R., Tu, S. and Reinberg, D. (2010). Molecular signals of epigenetic states. *Science* 330, 612-6.

14.Boyle, M., Wong, C., Rocha, M. and Jones, D. L. (2007). Decline in self-renewal factors contributes to aging of the stem cell niche in the Drosophila testis. *Cell Stem Cell* 1, 470-8.

15.Brawley, C. and Matunis, E. (2004). Regeneration of male germline stem cells by spermatogonial dedifferentiation in vivo. *Science* 304, 1331-4.

16.Buszczak, M., Paterno, S. and Spradling, A. C. (2009). Drosophila stem cells share a common requirement for the histone H2B ubiquitin protease scrawny. *Science* 323, 248-51.

17.Buszczak, M. and Spradling, A. C. (2006). Searching chromatin for stem cell identity. *Cell* 125, 233-6.

18.Casper, A. L., Baxter, K. and Van Doren, M. (2011). no child left behind encodes a novel chromatin factor required for germline stem cell maintenance in males but not females. *Development* 138, 3357-66.

19.Chen, C. C., Carson, J. J., Feser, J., Tamburini, B., Zabaronick, S., Linger, J. and Tyler, J. K. (2008). Acetylated lysine 56 on histone H3 drives chromatin assembly after repair and signals for the completion of repair. *Cell* 134, 231-43.

20.Chen, D. and McKearin, D. (2003a). Dpp signaling silences bam transcription directly to establish asymmetric divisions of germline stem cells. *Curr Biol* 13, 1786-91.

21.Chen, D. and McKearin, D. M. (2003b). A discrete transcriptional silencer in the bam gene determines asymmetric division of the Drosophila germline stem cell. *Development* 130, 1159-70.

22.Chen, H., Chen, X. and Zheng, Y. (2013). Nuclear lamina regulates the organization of the stem-cell niche. *Cell Stem Cell, in press.*

23.Chen, X. (2008). Stem cells: What can we learn from flies? *Fly (Austin)* 2.

24.Chen, X., Hiller, M., Sancak, Y. and Fuller, M. T. (2005). Tissue-specific TAFs counteract Polycomb to turn on terminal differentiation. *Science* 310, 869-72.

25.Chen, X., Lu, C., Prado, J. R., Eun, S. H. and Fuller, M. T. (2011). Sequential changes at differentiation gene promoters as they become active in a stem cell lineage. *Development* 138, 2441-50.

26.Cheng, J., Tiyaboonchai, A., Yamashita, Y. M. and Hunt, A. J. (2011). Asymmetric division of cyst stem cells in Drosophila testis is ensured by anaphase spindle repositioning. *Development* 138, 831-7.

27.Cheng, J., Turkel, N., Hemati, N., Fuller, M. T., Hunt, A. J. and Yamashita, Y. M. (2008). Centrosome misorientation reduces stem cell division during ageing. *Nature* 456, 599-604.

28.Cherry, C. M. and Matunis, E. L. (2010). Epigenetic regulation of stem cell maintenance in the Drosophila testis via the nucleosome-remodeling factor NURF. *Cell Stem Cell* 6, 557-67.

29.Cinalli, R. M., Rangan, P. and Lehmann, R. (2008). Germ cells are forever. *Cell* 132, 559-62.

30.Clarke, M. F. and Fuller, M. (2006). Stem cells and cancer: two faces of eve. *Cell* 124, 1111-5.

31.Cler, E., Papai, G., Schultz, P. and Davidson, I. (2009). Recent advances in understanding the structure and function of general transcription factor TFIID. *Cell Mol Life Sci* 66, 2123-34.

32.Clevers, H. (2005). Stem cells, asymmetric division and cancer. *Nat Genet* 37, 1027-8.

33.Corpet, A. and Almouzni, G. (2009). Making copies of chromatin: the challenge of nucleosomal organization and epigenetic information. *Trends Cell Biol* 19, 29-41.

34.Davies, E. L. and Fuller, M. T. (2008). Regulation of self-renewal and differentiation in adult stem cell lineages: lessons from the Drosophila male germ line. *Cold Spring Harb Symp Quant Biol* 73, 137-45.

35.Davies, E. L. and Fuller, M. T. (2009). Regulation of Self-renewal and Differentiation in Adult Stem Cell Lineages: Lessons from the Drosophila Male Germ Line. *Cold Spring Harb Symp Quant Biol*.

36.de Cuevas, M. and Matunis, E. L. (2011). The stem cell niche: lessons from the Drosophila testis. *Development* 138, 2861-9.

37.De Koning, L., Corpet, A., Haber, J. E. and Almouzni, G. (2007). Histone chaperones: an escort network regulating histone traffic. *Nat Struct Mol Biol* 14, 997-1007.

38.Deal, R. B., Henikoff, J. G. and Henikoff, S. (2010). Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science* 328, 1161-4.

39.Deng, X., Hiatt, J. B., Nguyen, D. K., Ercan, S., Sturgill, D., Hillier, L. W., Schlesinger, F., Davis, C. A., Reinke, V. J., Gingeras, T. R. et al. (2011). Evidence for compensatory upregulation of expressed X-linked genes in mammals, Caenorhabditis elegans and Drosophila melanogaster. *Nat Genet* 43, 1179-85.

40.Dion, M. F., Kaplan, T., Kim, M., Buratowski, S., Friedman, N. and Rando, O. J. (2007). Dynamics of replication-independent histone turnover in budding yeast. *Science* 315, 1405-8.

41.Doggett, K., Jiang, J., Aleti, G. and White-Cooper, H. (2011). Wake-up-call, a lin-52 paralogue, and Always early, a lin-9 homologue physically interact, but have opposing functions in regulating testis-specific gene expression. *Dev Biol* 355, 381-93.

42.Drane, P., Ouararhni, K., Depaux, A., Shuaib, M. and Hamiche, A. (2010). The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev* 24, 1253-65.

43.Eberhart, C. G., Maines, J. Z. and Wasserman, S. A. (1996). Meiotic cell cycle requirement for a fly homologue of human Deleted in Azoospermia. *Nature* 381, 783-5.

44.Eliazer, S., Shalaby, N. A. and Buszczak, M. (2011). Loss of lysine-specific demethylase 1 nonautonomously causes stem cell tumors in the Drosophila ovary. *Proc Natl Acad Sci U S A* 108, 7064-9.

45.Eun, S. H., Gan, Q. and Chen, X. (2010). Epigenetic regulation of germ cell differentiation. *Curr Opin Cell Biol* 22, 737-43.

46.Eun, S. H., Stoiber, P. M., Wright, H. J., McMurdie, K. E., Choi, C. H., Gan, Q., Lim, C. and Chen, X. (2013). MicroRNAs downregulate Bag of marbles to ensure proper terminal differentiation in the Drosophila male germline. *Development* 140, 23-30.

47.Fischle, W., Tseng, B. S., Dormann, H. L., Ueberheide, B. M., Garcia, B. A., Shabanowitz, J., Hunt, D. F., Funabiki, H. and Allis, C. D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* 438, 1116-22.

48.Fischle, W., Wang, Y. and Allis, C. D. (2003). Binary switches and modification cassettes in histone biology and beyond. *Nature* 425, 475-9.

49.Flaherty, M. S., Salis, P., Evans, C. J., Ekas, L. A., Marouf, A., Zavadil, J., Banerjee, U. and Bach, E. A. (2010). chinmo is a functional effector of the JAK/STAT pathway that regulates eye development, tumor formation, and stem cell self-renewal in Drosophila. *Dev Cell* 18, 556-68.

50.Friedmann-Morvinski, D., Bushong, E. A., Ke, E., Soda, Y., Marumoto, T., Singer, O., Ellisman, M. H. and Verma, I. M. (2012). Dedifferentiation of neurons and astrocytes by oncogenes can induce gliomas in mice. *Science* 338, 1080-4.

51.Fuller, M. T. (1998). Genetic control of cell proliferation and differentiation in Drosophila spermatogenesis. *Semin Cell Dev Biol* 9, 433-44.

52.Fuller, M. T. and Spradling, A. C. (2007). Male and female Drosophila germline stem cells: two versions of immortality. *Science* 316, 402-4.

53.Gan, Q., Chepelev, I., Wei, G., Tarayrah, L., Cui, K., Zhao, K. and Chen, X. (2010a). Dynamic regulation of alternative splicing and chromatin structure in Drosophila gonads revealed by RNA-seq. *Cell Res* 20, 763-83.

54.Gan, Q., Schones, D. E., Ho Eun, S., Wei, G., Cui, K., Zhao, K. and Chen, X. (2010b). Monovalent and unpoised status of most genes in undifferentiated cell-enriched Drosophila testis. *Genome Biol* 11, R42.

55.Gasser, R., Koller, T. and Sogo, J. M. (1996). The stability of nucleosomes at the replication fork. *J Mol Biol* 258, 224-39.

56.Gelbart, M. E., Larschan, E., Peng, S., Park, P. J. and Kuroda, M. I. (2009). Drosophila MSL complex globally acetylates H4K16 on the male X chromosome for dosage compensation. *Nat Struct Mol Biol* 16, 825-32.

57.Goldberg, A. D., Banaszynski, L. A., Noh, K. M., Lewis, P. W., Elsaesser, S. J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X. et al. (2010). Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* 140, 678-91.

58.Gonczy, P., Matunis, E. and DiNardo, S. (1997). bag-of-marbles and benign gonial cell neoplasm act in the germline to restrict proliferation during Drosophila spermatogenesis. *Development* 124, 4361-71.

59.Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. and Young, R. A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 130, 77-88.

60.Guilbaud, G., Rappailles, A., Baker, A., Chen, C. L., Arneodo, A., Goldar, A., d'Aubenton-Carafa, Y., Thermes, C., Audit, B. and Hyrien, O. (2011). Evidence for sequential and increasing activation of replication origins along replication timing gradients in the human genome. *PLoS Comput Biol* 7, e1002322.

61.Gupta, V., Parisi, M., Sturgill, D., Nuttall, R., Doctolero, M., Dudko, O. K., Malley, J. D., Eastman, P. S. and Oliver, B. (2006). Global analysis of X-chromosome dosage compensation. *J Biol* 5, 3.

62.Hales, K. G. and Fuller, M. T. (1997). Developmentally regulated mitochondrial fusion mediated by a conserved, novel, predicted GTPase. *Cell* 90, 121-9.

63.Henikoff, S., Furuyama, T. and Ahmad, K. (2004a). Histone variants, nucleosome assembly and epigenetic inheritance. *Trends Genet* 20, 320-6.

64.Henikoff, S., McKittrick, E. and Ahmad, K. (2004b). Epigenetics, histone H3 variants, and the inheritance of chromatin states. *Cold Spring Harb Symp Quant Biol* 69, 235-43.

65.Hiller, M., Chen, X., Pringle, M. J., Suchorolski, M., Sancak, Y., Viswanathan, S., Bolival, B., Lin, T. Y., Marino, S. and Fuller, M. T. (2004). Testis-specific TAF homologs collaborate to control a tissue-specific transcription program. *Development* 131, 5297-308.

66.Hiller, M. A., Lin, T. Y., Wood, C. and Fuller, M. T. (2001). Developmental regulation of transcription by a tissue-specific TAF homolog. *Genes Dev* 15, 1021-30.

67.Hime, G. R., Brill, J. A. and Fuller, M. T. (1996). Assembly of ring canals in the male germ line from structural components of the contractile ring. *J Cell Sci* 109 (Pt 12), 2779-88.

68.Hung, M. S., Karthikeyan, N., Huang, B., Koo, H. C., Kiger, J. and Shen, C. J. (1999). Drosophila proteins related to vertebrate DNA (5-cytosine) methyltransferases. *Proc Natl Acad Sci U S A* 96, 11940-5.

69.Inaba, M. and Yamashita, Y. M. (2012). Asymmetric stem cell division: precision for robustness. *Cell Stem Cell* 11, 461-9.

70.Inaba, M., Yuan, H., Salzmann, V., Fuller, M. T. and Yamashita, Y. M. (2010). E-cadherin is required for centrosome and spindle orientation in Drosophila male germline stem cells. *PLoS One* 5, e12473.

71.Insco, M. L., Bailey, A. S., Kim, J., Olivares, G. H., Wapinski, O. L., Tam, C. H. and Fuller, M. T. (2012). A self-limiting switch based on translational control regulates the

transition from proliferation to differentiation in an adult stem cell lineage. *Cell Stem Cell* 11, 689-700.

72.Insco, M. L., Leon, A., Tam, C. H., McKearin, D. M. and Fuller, M. T. (2009). Accumulation of a differentiation regulator specifies transit amplifying division number in an adult stem cell lineage. *Proc Natl Acad Sci U S A* 106, 22311-6.

73.Irvine, D. V., Zaratiegui, M., Tolia, N. H., Goto, D. B., Chitwood, D. H., Vaughn, M. W., Joshua-Tor, L. and Martienssen, R. A. (2006). Argonaute slicing is required for heterochromatic silencing and spreading. *Science* 313, 1134-7.

74.Issigonis, M., Tulina, N., de Cuevas, M., Brawley, C., Sandler, L. and Matunis, E. (2009). JAK-STAT signal inhibition regulates competition in the Drosophila testis stem cell niche. *Science* 326, 153-6.

75.Itman, C. and Loveland, K. L. (2008). SMAD expression in the testis: an insight into BMP regulation of spermatogenesis. *Dev Dyn* 237, 97-111.

76.Iwai, Y., Usui, T., Hirano, S., Steward, R., Takeichi, M. and Uemura, T. (1997). Axon patterning requires DN-cadherin, a novel neuronal adhesion receptor, in the Drosophila embryonic CNS. *Neuron* 19, 77-89.

77.Jackson, V. and Chalkley, R. (1981a). A new method for the isolation of replicative chromatin: selective deposition of histone on both new and old DNA. *Cell* 23, 121-34.

78.Jackson, V. and Chalkley, R. (1981b). A reevaluation of new histone deposition on replicating chromatin. *J Biol Chem* 256, 5095-103.

79.Jacobs, J. J. and van Lohuizen, M. (2002). Polycomb repression: from cellular memory to cellular proliferation and cancer. *Biochim Biophys Acta* 1602, 151-61.

80.Jaenisch, R. and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* 132, 567-82.

81.Jenuwein, T. and Allis, C. D. (2001). Translating the histone code. *Science* 293, 1074-80.
Jiang, J., Benson, E., Bausek, N., Doggett, K. and White-Cooper, H. (2007). Tombola, a tesmin/TSO1-family protein, regulates transcriptional activation in the Drosophila male germline and physically interacts with always early. *Development* 134, 1549-59.

82.Jiang, J. and White-Cooper, H. (2003). Transcriptional activation in Drosophila spermatogenesis involves the mutually dependent function of aly and a novel meiotic arrest gene cookie monster. *Development* 130, 563-73.

83.Jin, Z., Kirilly, D., Weng, C., Kawase, E., Song, X., Smith, S., Schwartz, J. and Xie, T. (2008). Differentiation-defective stem cells outcompete normal stem cells for niche occupancy in the Drosophila ovary. *Cell Stem Cell* 2, 39-49.

84.Kai, T. and Spradling, A. (2004). Differentiating germ cells can revert into functional stem cells in Drosophila melanogaster ovaries. *Nature* 428, 564-9.

85.Kamakaka, R. T. and Biggins, S. (2005). Histone variants: deviants? *Genes Dev* 19, 295-310.

86.Kawase, E., Wong, M. D., Ding, B. C. and Xie, T. (2004). Gbb/Bmp signaling is essential for maintaining germline stem cells and for repressing bam transcription in the Drosophila testis. *Development* 131, 1365-75.

87.Kiger, A. A., Jones, D. L., Schulz, C., Rogers, M. B. and Fuller, M. T. (2001). Stem cell self-renewal specified by JAK-STAT activation in response to a support cell cue. *Science* 294, 2542-5.

88.Kiger, A. A., White-Cooper, H. and Fuller, M. T. (2000). Somatic support cells restrict germline stem cell self-renewal and promote differentiation. *Nature* 407, 750-4.

89.Kleifeld, O., Doucet, A., auf dem Keller, U., Prudova, A., Schilling, O., Kainthan, R. K., Starr, A. E., Foster, L. J., Kizhakkedathu, J. N. and Overall, C. M. (2010). Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. *Nat Biotechnol* 28, 281-8.

90.Knoblich, J. A. (2008). Mechanisms of asymmetric stem cell division. *Cell* 132, 583-97.

91.Konev, A. Y., Tribus, M., Park, S. Y., Podhraski, V., Lim, C. Y., Emelyanov, A. V., Vershilova, E., Pirrotta, V., Kadonaga, J. T., Lusser, A. et al. (2007). CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin in vivo. *Science* 317, 1087-90.

92.Kouskouti, A. and Talianidis, I. (2005). Histone modifications defining active genes persist after transcriptional and mitotic inactivation. *EMBO J* 24, 347-57.

93.Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693-705.

94.Kurimoto, K., Yabuta, Y., Ohinata, Y., Ono, Y., Uno, K. D., Yamada, R. G., Ueda, H. R. and Saitou, M. (2006). An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. *Nucleic Acids Res* 34, e42.

95.Kurimoto, K., Yabuta, Y., Ohinata, Y. and Saitou, M. (2007). Global single-cell cDNA amplification to provide a template for representative high-density oligonucleotide microarray analysis. *Nat Protoc* 2, 739-52.

96.Le Bras, S. and Van Doren, M. (2006). Development of the male germline stem cell niche in Drosophila. *Dev Biol* 294, 92-103.

97.Leatherman, J. L. and Dinardo, S. (2008). Zfh-1 controls somatic stem cell self-renewal in the Drosophila testis and nonautonomously influences germline stem cell self-renewal. *Cell Stem Cell* 3, 44-54.

98.Leatherman, J. L. and Dinardo, S. (2010). Germline self-renewal requires cyst stem cells and stat regulates niche adhesion in Drosophila testes. *Nat Cell Biol* 12, 806-11.

99.Li, Q., Zhou, H., Wurtele, H., Davies, B., Horazdovsky, B., Verreault, A. and Zhang, Z. (2008). Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. *Cell* 134, 244-55.

100.Lim, C., Tarayrah, L. and Chen, X. (2012). Transcriptional regulation during Drosophila spermatogenesis. *Spermatogenesis* 2, 158-166.

101.Lim, J. G. and Fuller, M. T. (2012). Somatic cell lineage is required for differentiation and not maintenance of germline stem cells in Drosophila testes. *Proc Natl Acad Sci U S A* 109, 18477-81.

102.Lin, T. Y., Viswanathan, S., Wood, C., Wilson, P. G., Wolf, N. and Fuller, M. T. (1996). Coordinate developmental control of the meiotic cell cycle and spermatid differentiation in Drosophila males. *Development* 122, 1331-41.
Losick, V. P., Morris, L. X., Fox, D. T. and Spradling, A. (2011). Drosophila stem cell niches: a decade of discovery suggests a unified view of stem cell regulation. *Dev Cell* 21, 159-71.

103.Luciani, N., Marie-Claire, C., Ruffet, E., Beaumont, A., Roques, B. P. and Fournie-Zaluski, M. C. (1998). Characterization of Glu350 as a critical residue involved in the N-terminal amine binding site of aminopeptidase N (EC 3.4.11.2): insights into its mechanism of action. *Biochemistry* 37, 686-92.

104.Lyko, F., Ramsahoye, B. H. and Jaenisch, R. (2000a). DNA methylation in Drosophila melanogaster. *Nature* 408, 538-40.

105.Lyko, F., Whittaker, A. J., Orr-Weaver, T. L. and Jaenisch, R. (2000b). The putative Drosophila methyltransferase gene dDnmt2 is contained in a transposon-like element and is expressed specifically in ovaries. *Mech Dev* 95, 215-7.

106.Manseau, L., Baradaran, A., Brower, D., Budhu, A., Elefant, F., Phan, H., Philp, A. V., Yang, M., Glover, D., Kaiser, K. et al. (1997). GAL4 enhancer traps expressed in the embryo, larval brain, imaginal discs, and ovary of Drosophila. *Dev Dyn* 209, 310-22.

107.Martin, C. and Zhang, Y. (2007). Mechanisms of epigenetic inheritance. *Curr Opin Cell Biol* 19, 266-72.

108.Masumoto, H., Hawke, D., Kobayashi, R. and Verreault, A. (2005). A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature* 436, 294-8.

109.Matangkasombut, O., Auty, R. and Buratowski, S. (2004). Structure and function of the TFIID complex. *Adv Protein Chem* 67, 67-92.

110.McKearin, D. M. and Spradling, A. C. (1990). bag-of-marbles: a Drosophila gene required to initiate both male and female gametogenesis. *Genes Dev* 4, 2242-51.

111.Metcalf, C. E. and Wassarman, D. A. (2007). Nucleolar colocalization of TAF1 and testis-specific TAFs during Drosophila spermatogenesis. *Dev Dyn* 236, 2836-43.

112.Mito, Y., Henikoff, J. G. and Henikoff, S. (2007). Histone replacement marks the boundaries of cis-regulatory domains. *Science* 315, 1408-11.

113.Moggs, J. G., Grandi, P., Quivy, J. P., Jonsson, Z. O., Hubscher, U., Becker, P. B. and Almouzni, G. (2000). A CAF-1-PCNA-mediated chromatin assembly pathway triggered by sensing DNA damage. ***Mol Cell Biol*** **20, 1206-18.**

114.Monk, A. C., Siddall, N. A., Volk, T., Fraser, B., Quinn, L. M., McLaughlin, E. A. and Hime, G. R. (2010). HOW is required for stem cell maintenance in the Drosophila testis and for the onset of transit-amplifying divisions. ***Cell Stem Cell*** **6, 348-60.**

115.Moon, S., Cho, B., Min, S. H., Lee, D. and Chung, Y. D. (2011). The THO complex is required for nucleolar integrity in Drosophila spermatocytes. ***Development*** **138, 3835-45.**

116.Morrison, S. J. and Kimble, J. (2006). Asymmetric and symmetric stem-cell divisions in development and cancer. ***Nature*** **441, 1068-74.**

117.Morrison, S. J. and Spradling, A. C. (2008). Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. ***Cell*** **132, 598-611.**

118.Motamedi, M. R., Verdel, A., Colmenares, S. U., Gerber, S. A., Gygi, S. P. and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. ***Cell*** **119, 789-802.**

119.Moustakas, A. and Heldin, C. H. (2009). The regulation of TGFbeta signal transduction. ***Development*** **136, 3699-714.**

120.Mulinari, S., Hacker, U. and Castillejo-Lopez, C. (2006). Expression and regulation of Spatzle-processing enzyme in Drosophila. ***FEBS Lett*** **580, 5406-10.**

121.Murali, T., Pacifico, S., Yu, J., Guest, S., Roberts, G. G., 3rd and Finley, R. L., Jr. (2011). DroID 2011: a comprehensive, integrated resource for protein, transcription factor, RNA and gene interactions for Drosophila. ***Nucleic Acids Res*** **39, D736-43.**

122., T., Sharma, M., Nabeshima, Y., Braun, R. E. and Yoshida, S. (2010). Functional hierarchy and reversibility within the murine spermatogenic stem cell compartment. ***Science*** **328, 62-7.**

123.Oatley, J. M. (2010). Spermatogonial stem cell biology in the bull: development of isolation, culture, and transplantation methodologies and their potential impacts on cattle production. ***Soc Reprod Fertil Suppl*** **67, 133-43.**

124.Oda, H. and Tsukita, S. (1999). Nonchordate classic cadherins have a structurally and functionally unique domain that is absent from chordate classic cadherins. ***Dev Biol*** **216, 406-22.**

125.Parisi, M., Nuttall, R., Naiman, D., Bouffard, G., Malley, J., Andrews, J., Eastman, S. and Oliver, B. (2003). Paucity of genes on the Drosophila X chromosome showing male-biased expression. ***Science*** **299, 697-700.**

126.Parrott, B. B., Chiang, Y., Hudson, A., Sarkar, A., Guichet, A. and Schulz, C. (2011). Nucleoporin98-96 function is required for transit amplification divisions in the germ line of Drosophila melanogaster. ***PLoS One*** **6, e25087.**

127.Parrott, B. B., Hudson, A., Brady, R. and Schulz, C. (2012). Control of germline stem cell division frequency--a novel, developmentally regulated role for epidermal growth factor signaling. *PLoS One* 7, e36460.

128.Perezgasga, L., Jiang, J., Bolival, B., Jr., Hiller, M., Benson, E., Fuller, M. T. and White-Cooper, H. (2004). Regulation of transcription of meiotic cell cycle and terminal differentiation genes by the testis-specific Zn-finger protein matotopetli. *Development* 131, 1691-702.

129.Rangan, P., DeGennaro, M. and Lehmann, R. (2008). Regulating gene expression in the Drosophila germ line. *Cold Spring Harb Symp Quant Biol* 73, 1-8.

130.Ray-Gallet, D., Quivy, J. P., Scamps, C., Martini, E. M., Lipinski, M. and Almouzni, G. (2002). HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. *Mol Cell* 9, 1091-100.

131.Recht, J., Tsubota, T., Tanny, J. C., Diaz, R. L., Berger, J. M., Zhang, X., Garcia, B. A., Shabanowitz, J., Burlingame, A. L., Hunt, D. F. et al. (2006). Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. *Proc Natl Acad Sci U S A* 103, 6988-93.

132.Richards, E. J. and Elgin, S. C. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* 108, 489-500.

133.Ringrose, L. and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu Rev Genet* 38, 413-43.

134.Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139-40.

135.Rorth, P. (1998). Gal4 in the Drosophila female germline. *Mech Dev* 78, 113-8.

136.Rushlow, C., Colosimo, P. F., Lin, M. C., Xu, M. and Kirov, N. (2001). Transcriptional regulation of the Drosophila gene zen by competing Smad and Brinker inputs. *Genes Dev* 15, 340-51.

137.Russev, G. and Hancock, R. (1981). Formation of hybrid nucleosomes cantaining new and old histones. *Nucleic Acids Res* 9, 4129-37.

138.Saha, A., Wittmeyer, J. and Cairns, B. R. (2006). Chromatin remodelling: the industrial revolution of DNA around histones. *Nat Rev Mol Cell Biol* 7, 437-47.

139.Sanson, B., White, P. and Vincent, J. P. (1996). Uncoupling cadherin-based adhesion from wingless signalling in Drosophila. *Nature* 383, 627-30.

140.Santel, A., Blumer, N., Kampfer, M. and Renkawitz-Pohl, R. (1998). Flagellar mitochondrial association of the male-specific Don Juan protein in Drosophila spermatozoa. *J Cell Sci* 111 (Pt 22), 3299-309.

141.Santel, A., Winhauer, T., Blumer, N. and Renkawitz-Pohl, R. (1997). The Drosophila don juan (dj) gene encodes a novel sperm specific protein component characterized by an unusual domain of a repetitive amino acid motif. ***Mech Dev*** **64, 19-30.**

142.Sarkar, A., Parikh, N., Hearn, S. A., Fuller, M. T., Tazuke, S. I. and Schulz, C. (2007). Antagonistic roles of Rac and Rho in organizing the germ cell microenvironment. ***Curr Biol*** **17, 1253-8.**

143.Schafer, M., Borsch, D., Hulster, A. and Schafer, U. (1993). Expression of a gene duplication encoding conserved sperm tail proteins is translationally regulated in Drosophila melanogaster. ***Mol Cell Biol*** **13, 1708-18.**

144.Schreiber, S. L. and Bernstein, B. E. (2002). Signaling network model of chromatin. ***Cell*** **111, 771-8.**

145.Schulz, C., Kiger, A. A., Tazuke, S. I., Yamashita, Y. M., Pantalena-Filho, L. C., Jones, D. L., Wood, C. G. and Fuller, M. T. (2004). A misexpression screen reveals effects of bag-of-marbles and TGF beta class signaling on the Drosophila male germ-line stem cell lineage. ***Genetics*** **167, 707-23.**

146.Schulz, C., Wood, C. G., Jones, D. L., Tazuke, S. I. and Fuller, M. T. (2002). Signaling from germ cells mediated by the rhomboid homolog stet organizes encapsulation by somatic support cells. ***Development*** **129, 4523-34.**

147.Schwartz, B. E. and Ahmad, K. (2005). Transcriptional activation triggers deposition and removal of the histone variant H3.3. ***Genes Dev*** **19, 804-14.**

148.Schwitalla, S., Fingerle, A. A., Cammareri, P., Nebelsiek, T., Goktuna, S. I., Ziegler, P. K., Canli, O., Heijmans, J., Huels, D. J., Moreaux, G. et al. (2013). Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. ***Cell*** **152, 25-38.**

149.Seydoux, G. and Braun, R. E. (2006). Pathway to totipotency: lessons from germ cells. ***Cell*** **127, 891-904.**

150.Sheng, X. R., Brawley, C. M. and Matunis, E. L. (2009). Dedifferentiating spermatogonia outcompete somatic stem cells for niche occupancy in the Drosophila testis. ***Cell Stem Cell*** **5, 191-203.**

151.Sheng, X. R. and Matunis, E. (2011). Live imaging of the Drosophila spermatogonial stem cell niche reveals novel mechanisms regulating germline stem cell output. ***Development*** **138, 3367-76.**

152.Sheng, X. R. and Matunis, E. L. (2009). Make room for dedifferentiation. ***Fly (Austin)*** **3, 283-5.**

153.Shibahara, K. and Stillman, B. (1999). Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. ***Cell*** **96, 575-85.**

154.Shivdasani, A. A. and Ingham, P. W. (2003). Regulation of stem cell maintenance and transit amplifying cell proliferation by tgf-beta signaling in Drosophila spermatogenesis. ***Curr Biol*** **13, 2065-72.**

155.Siddall, N. A., McLaughlin, E. A., Marriner, N. L. and Hime, G. R. (2006). The RNA-binding protein Musashi is required intrinsically to maintain stem cell identity. *Proc Natl Acad Sci U S A* 103, 8402-7.

156.Silver, D. L., Geisbrecht, E. R. and Montell, D. J. (2005). Requirement for JAK/STAT signaling throughout border cell migration in Drosophila. *Development* 132, 3483-92.
Smith, D. J. and Whitehouse, I. (2012). Intrinsic coupling of lagging-strand synthesis to chromatin assembly. *Nature* 483, 434-8.

157.Song, X., Wong, M. D., Kawase, E., Xi, R., Ding, B. C., McCarthy, J. J. and Xie, T. (2004). Bmp signals from niche cells directly repress transcription of a differentiation-promoting gene, bag of marbles, in germline stem cells in the Drosophila ovary. *Development* 131, 1353-64.

158.Szenker, E., Ray-Gallet, D. and Almouzni, G. (2011). The double face of the histone variant H3.3. *Cell Res* 21, 421-34.

159.Tagami, H., Ray-Gallet, D., Almouzni, G. and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* 116, 51-61.

160.Talchai, C., Xuan, S., Lin, H. V., Sussel, L. and Accili, D. (2012). Pancreatic beta cell dedifferentiation as a mechanism of diabetic beta cell failure. *Cell* 150, 1223-34.

161.Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui, A. et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* 6, 377-82.

162.Tarayrah, L., Herz, H. M., Shilatifard, A. and Chen, X. (2013). Histone demethylase dUTX antagonizes JAK-STAT signaling to maintain proper gene expression and architecture of the Drosophila testis niche. *Development* 140, 1014-23.

163.Terry, N. A., Tulina, N., Matunis, E. and DiNardo, S. (2006). Novel regulators revealed by profiling Drosophila testis stem cells within their niche. *Dev Biol* 294, 246-57.

164.Toledano, H., D'Alterio, C., Czech, B., Levine, E. and Jones, D. L. (2012). The let-7-Imp axis regulates ageing of the Drosophila testis stem-cell niche. *Nature* 485, 605-10.

165.Tora, L. (2002). A unified nomenclature for TATA box binding protein (TBP)-associated factors (TAFs) involved in RNA polymerase II transcription. *Genes Dev* 16, 673-5.

166.Tran, J., Brenner, T. J. and DiNardo, S. (2000). Somatic control over the germline stem cell lineage during Drosophila spermatogenesis. *Nature* 407, 754-7.

167.Tran, V., Lim, C., Xie, J. and Chen, X. (2012). Asymmetric division of Drosophila male germline stem cell shows asymmetric histone distribution. *Science* 338, 679-82.

168.Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J. and Pachter, L. (2010). Transcript assembly and quantification

by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. ***Nat Biotechnol*** **28, 511-5.**

169.Tulina, N. and Matunis, E. (2001). Control of stem cell self-renewal in Drosophila spermatogenesis by JAK-STAT signaling. ***Science*** **294, 2546-9.**

170.Turner, B. M. (2002). Cellular memory and the histone code. ***Cell*** **111, 285-91.**

171.Valls, E., Sanchez-Molina, S. and Martinez-Balbas, M. A. (2005). Role of histone modifications in marking and activating genes through mitosis. ***J Biol Chem*** **280, 42592-600.**

172.Van Doren, M., Williamson, A. L. and Lehmann, R. (1998). Regulation of zygotic gene expression in Drosophila primordial germ cells. ***Curr Biol*** **8, 243-6.**

173.Verreault, A., Kaufman, P. D., Kobayashi, R. and Stillman, B. (1996). Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. ***Cell*** **87, 95-104.**

174.Wallenfang, M. R., Nayak, R. and DiNardo, S. (2006). Dynamics of the male germline stem cell population during aging of Drosophila melanogaster. ***Aging Cell*** **5, 297-304.**

175.Wang, Y., Wang, L. and Wang, Z. (2008). Transgenic analyses of TGIF family proteins in Drosophila imply their role in cell growth. ***J Genet Genomics*** **35, 457-65.**

176.Wang, Z., Gerstein, M. and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. ***Nat Rev Genet*** **10, 57-63.**

177.Wang, Z. and Mann, R. S. (2003). Requirement for two nearly identical TGIF-related homeobox genes in Drosophila spermatogenesis. ***Development*** **130, 2853-65.**

178.White-Cooper, H. (2009). Studying how flies make sperm--investigating gene function in Drosophila testes. ***Mol Cell Endocrinol*** **306, 66-74.**

179.White-Cooper, H., Leroy, D., MacQueen, A. and Fuller, M. T. (2000). Transcription of meiotic cell cycle and terminal differentiation genes depends on a conserved chromatin associated protein, whose nuclear localisation is regulated. ***Development*** **127, 5463-73.**

180.White-Cooper, H., Schafer, M. A., Alphey, L. S. and Fuller, M. T. (1998). Transcriptional and post-transcriptional control mechanisms coordinate the onset of spermatid differentiation with meiosis I in Drosophila. ***Development*** **125, 125-34.**

181.Wong, C. and Jones, D. L. (2012). Efficiency of spermatogonial dedifferentiation during aging. ***PLoS One*** **7, e33635.**

182.Xu, M., Long, C., Chen, X., Huang, C., Chen, S. and Zhu, B. (2010). Partitioning of histone H3-H4 tetramers during DNA replication-dependent chromatin assembly. ***Science*** **328, 94-8.**

183.Yadlapalli, S., Cheng, J. and Yamashita, Y. M. (2011). Drosophila male germline stem cells do not asymmetrically segregate chromosome strands. ***J Cell Sci*** **124, 933-9.**

184.Yadlapalli, S. and Yamashita, Y. M. (2013). Chromosome-specific nonrandom sister chromatid segregation during stem-cell division. ***Nature*****.**

185.Yamashita, Y. M. (2010). Cell adhesion in regulation of asymmetric stem cell division. ***Curr Opin Cell Biol*** **22, 605-10.**

186.Yamashita, Y. M., Jones, D. L. and Fuller, M. T. (2003). Orientation of asymmetric stem cell division by the APC tumor suppressor and centrosome. ***Science*** **301, 1547-50.**

187.Yamashita, Y. M., Mahowald, A. P., Perlin, J. R. and Fuller, M. T. (2007). Asymmetric inheritance of mother versus daughter centrosome in stem cell division. ***Science*** **315, 518-21.**

188.Yamashita, Y. M., Yuan, H., Cheng, J. and Hunt, A. J. (2010). Polarity in stem cell division: asymmetric stem cell division in tissue homeostasis. ***Cold Spring Harb Perspect Biol*** **2, a001313.**

189.Yang, S. Y., Baxter, E. M. and Van Doren, M. (2012). Phf7 controls male sex determination in the Drosophila germline. ***Dev Cell*** **22, 1041-51.**

190.Zhang, H., Tan, J., Reynolds, E., Kuebler, D., Faulhaber, S. and Tanouye, M. (2002). The Drosophila slamdance gene: a mutation in an aminopeptidase can cause seizure, paralysis and neuronal failure. ***Genetics*** **162, 1283-99.**

Cindy Lim

Education

2002-2005	B.S.Biochemistry/Molecular Biology	University of Georgia, Athens, GA
2007- present	Ph. D. Biology	Johns Hopkins University, Baltimore, MD

Personal Statement

I am a trained Development Biologist and RNA Biologist. My area of expertise is Stem Cells Biology, Biochemistry, Drosophila Genetics and Transgene generation, RNA Biology, miRNA library generation and detection in tissue samples using *in situ hybridization*, mRNA sequencing from single cysts, Epigenetic modifications in histones, Chromatin Immunoprecipitation, and Enzymology of metalloproteinase. My overall research interest is to elucidate how the niche controls stems cell behavior, through potential factors such as epigenetic markers or proteins that regulate the extracellular matrix, currently focusing on male stem cells niche in *Drosophila*. Proliferating cells have been shown to repopulate the niche during aging to maintain tissue homeostasis. I am interested in studying how the stem cells niche functions to retain the stem cells and how proliferating cells repopulate the niche. Understanding the mechanism and factors involved in these processes offers the opportunity to establish new methods in tissue rejuvenation and disease treatment.

Professional Experience

2003-2004	Laboratory Technician with Dr. Wyatt Anderson, Department of Genetics, University of Georgia
2004	Undergraduate Research Assistant with Dr. Russell Malmberg, Department of Plant Biology, University of Georgia
2004-2005	Undergraduate Research Assistant with Dr. Michael Adams, Department of Biochemistry and Molecular Biology, University of Georgia
2005-2007	Research Assistant with Dr. Thomas Tuschl, RNA Biology, Rockefeller University
2007-present	Graduate Research Assistant with Dr. Xin Chen, Department of Biology, Johns Hopkins University
2013	Internship with Dr. Xavier de Mollerat du Jeu, R&D – Transfection and Molecular Delivery Division, Life Technologies

Teaching Experience

2008	Laboratory Teaching Assistant for Biochemistry, Johns Hopkins University
2009	Laboratory Teaching Assistant for Developmental Biology, Johns Hopkins University
2011	Teaching Assistant for Genetics, Johns Hopkins University
2012	Teaching Assistant for Developmental Biology, Johns Hopkins University

Awards and Honors

2002	Dean's List	University of Georgia
2003-2005	Presidential Scholar	University of Georgia
2005	Summa Cum Laude with Honors	University of Georgia
2011	Travel Grant from International Proteolysis Society	
2011	Outstanding Poster Award at Gordon Research Conference	
2011	Speaker at 7th General Meeting of International Proteolysis Society	

Professional Memberships

2005-2007	New York Academy of Science
2005-2007	Oligonucleotide Therapeutics Society
2007--present	Genetics Society of America
2010-2011	Society for Developmental Biology
2010-2012	International Proteolysis Society

Activities

2002-2005	International Student Ambassador	University of Georgia
2003-2005	President of Indonesian Student Organization	University of Georgia
2004-2005	Volunteer for Athens Clarke County Middle School	University of Georgia
2004-2005	Tutor for Chemistry and Biology	University of Georgia
2008-2009	Laboratory Teaching Assistant, Biochemistry	Johns Hopkins Univ.
2009-present	Volunteer, Teaching science in Baltimore High Schools	Johns Hopkins Univ.

2009	Laboratory Teaching Assistant, Developmental Biology	Johns Hopkins Univ.
2011	Teaching Assistant, Genetics	Johns Hopkins Univ.
2012	Teaching Assistant, Developmental Biology	Johns Hopkins Univ.
2012-present	Secretary, Indonesian Club	Johns Hopkins Univ.

Publications

I. Thesis

Ph.D. thesis: Analysis of Transcriptional and Epigenetic Regulation in *Drosophila melanogaster* Male Germ Cells Differentiation

II. Reviews

1. **Lim, C**.*, Tarayrah, L.*, Chen, X., (* equal contribution) "Transcriptional regulation during Drosophila spermatogenesis." Spermatogenesis, 2(3): 158-166. (2012)

III. Research Publications

1. **Lim, C**. and Chen, X., "A novel niche-specific aminopeptidase regulates dedifferentiation of progenitor germ cells in Drosophila adult testis" *Reviewed by Science and in preparation for re-submission.*

2**. Lim, C**., Shi, Z., Mejia, R., Mo, A., Tran, V., Zhao, K., Chen, X., "High resolution transcriptome analysis of single germline cysts in Drosophila wild type testis" *Manuscript in preparation*.

3. Tran, V*., **Lim, C***,. Xie, J., Chen, X., (* equal contribution) "Asymmetric histone inheritance in Drosophila male germline stem cells." Science,338(6107): 679-82 . (2012)

4. Eun, S., Stoiber, P., Wright, H., McMurdie, K., Choi, C., Gan, Q,. **Lim, C**., Chen, X., "MicroRNAs downregulate bag of marbles to ensure proper terminal differentiation in Drosophila male germ cells lineage." *Accepted by Development*. (2012)

5. Pena, J. T., Sohn-Lee, C., . Rouhanifard, S. H, Ludwig, J., Hafner, M., Mihailovic, A**., Lim, C**,. Holoch, D., Berninger, M., Zavolan, M., Tuschl, T., " miRNA in situ hybridization in formaldehyde and EDC-fixed tissues" Nat. Methods, 6: 139-141. (2009)

6. Hafner, M., Landgraf, p., Ludwig, J., Rice, A., Ojo, T., Holoch, D**., Lim, C**., Tuschl, T., "Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing" Methods, 44: 3-12. (2008)

Invited Presentations

1.Slamdance, an aminopeptidase required for niche architecture and stem cells maintenance. 7th General Meeting of International Proteolysis Society, San Diego 2011

Abstracts/Poster Presentations

1.Slamdance encodes a hub exclusive aminopeptidase that maintains stem cells in male Drosophila testis. 54th Drosophila Research Conference, Washington DC, 2013

2. Asymmetric division of Drosophila male germline stem cells shows asymmetric histone distribution. Cold Spring Harbor Laboratory – Base Pair to Body Plan Celebrating 60 Years of DNA, 2013

3. Slamdance, an aminopeptidase required for niche architecture and stem cells maintenance. 7th General Meeting of the International Proteolysis Society, 2011

4. Slamdance, an aminopeptidase required for niche architecture and stem cells maintenance. Gordon Research Conference: Regulated Proteolysis of Cell Surface Proteins, 2011

5. Slamdance is required for niche function in Drosophila testis. 51st Annual Drosophila Research Conference, 2010

6. High resolution transcriptome analysis of Drosophila male germline stem cells. Mid-Atlantic Society for Developmental Biology meeting, 2010

7. A high resolution transcriptome analysis of Drosophila male germline stem cells. 50th Annual Drosophila Research Conference, 2009